SEQUENCE LISTING

```
<110> Gentide Biopharmaceuticals, Inc.
     Bussell, Stuart
<120> METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENCES AS
DIRECT FUSIONS OR WITH LINKERS
<130>
      GNT-00101.P.1-US
      US 60/396,466
<150>
<151>
      2002-07-16
<160> 86
<170> PatentIn version 3.0
<210>
<211>
      573
<212> DNA
<213> Homo sapiens
<400> 1
ttcccaacca ttcccttatc caggettttt gacaacgeta tgctccgcgc ccatcgtctg
                                                                       60
caccagetgg cetttgacae ctaccaggag tttgaagaag cetatateee aaaggaacag
                                                                      120
aagtatteat teetgeagaa ceeccagaee teeetetgtt teteagagte tatteegaea
                                                                      180
ccctccaaca gggaggaaac acaacagaaa tccaacctag agctgctccg catctccctg
                                                                      240
ctgctcatcc agtcgtggct ggagcccgtg cagttcctca ggagtgtctt cgccaacagc
                                                                     300
ctggtgtacg gcgcctctga cagcaacgtc tatgacctcc taaaggacct agaggaaggc
                                                                      360
atccaaacgc tgatggggag gctggaagat ggcagcccc ggactgggca gatcttcaag
                                                                      420
cagacctaca gcaagttcga cacaaactca cacaacgatg acgcactact caagaactac
                                                                     480
gggctgctct actgcttcag gaaggacatg gacaaggtcg agacattcct gcgcatcgtg
                                                                      540
                                                                      573
cagtgccgct ctgtggaggg cagctgtggc ttc
<210>
       2
<211>
      191
<212>
      PRT
<213> Homo sapiens
<220>
<221>
      mat_peptide
<222>
      (1)..()
<400>
```

Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg

| Ala | HIS | Arg | 20 | HIS | GIII | ьец | Ala | 25 | Asp | 1111 | TÀT | GIII | 30 | Pne | GIU | |
|--------------------------|------------|-------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|----|
| Glu | Ala | Tyr 35 | Ile | Pro | Lys | Glu | Gln 40 | Lys | Tyr | Ser | Phe | Leu 45 | Gln | Asn | Pro | |
| Gln | Thr 50 | Ser | Leu | Cys | Phe | Ser 55 | Glu | Ser | Ile | Pŗo | Thr 60 | Pro | Ser | Asn | Arg | |
| Glu 65 | Glu | Thr | Gln | Gln | Lys 70 | Ser | Asn | Leu | Glu | Leu 75 | Leu | Arg | Ile | Ser | Leu 80 | |
| Leu | Leu | Ile | Gln | Ser 85 | Trp | Leu | Glu | Pro | Val 90 | Gļn | Phe | Leu | Arg | Ser 95 | Val | |
| Phe | Ala | Asn | Ser 100 | Leu | Val | Tyr | Gly | Ala 105 | Ser | Asp | Ser | Asn | Val 110 | Tyr | Asp | |
| Leu | Leu | Lys 115 | Asp | Leu | Glu | Glu | Gly 120 | Ile | Gln | Thr | Leu | Met 125 | Gly | Arg | Leu | |
| | Asp 130 | Gly | Ser | Pro | Arg | Thr 135 | Gly | Gln | Ile | Phe | Lys 140 | Gln | Thr | Tyr | Ser | |
| Lys 145 | Phe | Asp | Thr | Asn | Ser 150 | His | Asn | Asp | Asp | Ala 155 | Leu | Leu | Lys | Asn | Tyr 160 | |
| Gly | Leu | Leu | Tyr | Cys 165 | Phe | Arg | Lys | Asp | Met 170 | Asp | Lys | Val | Glu | Thr 175 | Phe | |
| | | Ile | Val 180 | Gln | Cys | Arg | Ser | Val 185 | Glu | Gly | Ser | Cys | Gly 190 | Phe | | |
| <21 <21 <21 <21 | 0 > | 3 38 DNA Artii | ficia | al | | | | | • | | | | | | | |
| <22 <22 | | syntl | netio | c se | quen | ce | | | ٠ | | | | | | | |
| <40 gga | | 3 cat a | atgti | tece | aa c | catt | ccct | t at | ccag | gc | | | | | | 38 |
| <21 <21 <21 <21 | 1> 2> | 4 36 DNA Arti: | ficia | al | | | | | | | | | | | | |
| <22 <22 | | syntl | heti | c se | quen | ce | | | | | | | | | | |
| <40 cgc | | 4 ccc 1 | taga | agcc | ac a | gctg | ccct | c ca | caga | | | | | | | 36 |

| <pre> <210> 6 <211> 21 <212> DNA <213> Artificial <220> <223> synthetic sequence <400> 6 tgctagttat tgctcagcgg t </pre> <pre> <pre> <210> 7 <211> 588 <212> DNA <213> Artificial </pre> <pre> <220> <221> 588 <212> DNA <213> Artificial </pre> <pre> <220> <223> synthetic sequence <400> 7 catatytic caaccattce cttatccagg ctttttgaca acgctatget ccgcgccat cgtctgcacc agctggctt tgacacctac caggagtttg aagaagccta tatcccaaag gaacagaagt attcattcct gcagaaccce cagacctcce tctgtttctc agagtctatt ccgacaccct ccaacagga ggaaacacaa cagaaatcca acctagaget gctccgcatc tccctgctgc tcatccagtc gtggctggag cccgtgcagt tcctcaggag tgtcttcgcc aacagcctgg tgtacggcg ctctgacage aacgtctatg acctcctaaa ggacctagag gaaggcatcc aaacgctgat ggggaggctg gaagatggca gccccggac tgggcagat ttcaagcaga cctacagcaa gttcgacaa aactcacaca acgatgacca actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgag aactacggc tgctctactg cttcaggaag gacatggaca aggtcgagaca aggtcgagaca attcctgag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgaagaca aggtcgaaca aggtcgaaca aggtcgaacaa aggtcgaacaa aggtcgaacaa aggtcgaacaa aggtcgaacaa aggtcgaacaa aggtcgaacaaacaaa aggacacaaaacaaa</pre></pre> | <210><211><212><213> | 5 19 DNA Artificial | | | | | |
|---|----------------------|------------------------------|------------|------------|------------|------------|-----|
| <pre>c210> 6 c211> 21 c212> DNA c213> Artificial c220> c223> synthetic sequence <400> 6 tgctagttat tgctcagcgg t c210> T c211> 588 c212> DNA c213> Artificial c220> c223> synthetic sequence c400> 6 tgctagttat tgctcagcgg t c210> T c211> 588 c212> DNA c213> Artificial c220> c223> synthetic sequence c400> T catatgttcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgcccat cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt ccgaccaccct ccaacaggga ggaaacacaa cagaaatcca acctagagct gctccgcatc tccctgctgc tcatccagtc gtggctggag cccgtgcagt tcctcaggag tgtcttcgcc aacagcctgg tgtacggcgc ctctgacagc aacgtcatg acctcctaaa ggacctagag gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc ttcaagcaga cctacagcaa gttcgacca aactcacaca acgatgacca actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctagag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagaca aggtcgagaca attcctgcgc</pre> | | synthetic seque | ence | | | | |
| <pre><211> 21 <212> DNA <213> Artificial </pre> <pre><220> <223> synthetic sequence <400> 6 tgctagttat tgctcagcgg t <210> 7 <211> 588 <212> DNA <213> Artificial </pre> <pre><220> <223> synthetic sequence </pre> <pre><400> 7 catatgttcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgccat cgtctgcacc agetggctt tgacacctac caggagtttg aagaagccta tatcccaaag gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt ccgacaccct ccaacaggag ggaaacacaa cagaaatcca acctagagct gctccgcatc tccctgctgc tcatccagtc gtggctggag cccgtgcagt tcctcagagg tgtcttcgcc aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacg actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54</pre> | | | | | | | 19 |
| <pre><223> synthetic sequence <4400> 6 tgctagttat tgctcagcgg t <210> 7 <211> 588 <212> DNA <213> Artificial <220> <223> synthetic sequence <4400> 7 catatgttcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgcccat cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag gaacagaagt attcattcct gcagaacccc cagacctcc tctgtttctc agagtctatt ccgacaccct ccaacaggag agaaacacaa cagaaatcca acctagagct gctccgcatc tccctgctgc tcatccagtc gtggctggag cccgtgcagt tcctcaggag tgtcttcgcc aacagcctgg tgtacggcgc ctctgacagc aacgtcatag acctcctaaa ggacctagag gaaggcatcc aaacgctgat ggggaggctg gaagatggca gccccggac tgggcagatc ttcaagcaga cctacagcaa gttcgacaa aactcacaca acgatagacg actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc</pre> | <211> <212> | 21 DNA | · | | | | |
| tgctagttat tgctcagcgg t <pre> <210> 7 <211> 588 <212> DNA <213> Artificial <220> <223> synthetic sequence <400> 7 catatgttcc caaccattcc cttatccagg ctttttgaca acgctatget ccgcgcccat cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt ccgacaccct ccaacaggga ggaaacacaa cagaaatcca acctagagct gctccgcatc tccctgctgc tcatccagtc gtggctggag cccgtgcagt tcctcaggag tgtcttcgcc aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacg actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 </pre> | | synthetic seque | ence | | | | |
| <pre><211> 588 <212> DNA <213> Artificial <220> <223> synthetic sequence <400> 7 catatgttcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgcccat cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt ccgacaccct ccaacaggga ggaaacacaa cagaaatcca acctagagct gctccgcatc tccctgctgc tcatccagtc gtggctggag cccgtgcagt tcctcaggag tgtcttcgcc aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag gaaggcatcc aaacgctgat ggggaggctg gaagatggca gccccggac tgggcagatc ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacg actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54</pre> | . • | _ | t | | | | 21 |
| <pre><223> synthetic sequence <400> 7 catatgttcc caaccattcc cttatccagg ctttttgaca acgetatget cegegeceat cgtctgcacc agetggeett tgacacctae caggagtttg aagaagecta tateccaaag gaacagaagt atteatteet geagaacece cagaceteee tetgtttete agagtetatt cegacaccet ccaacaggga ggaaacacaa cagaaateca acetagaget geteegeate tecetgetge teatecagte gtggetggag eeegtgeagt teeteaggag tgtettegee aacageetgg tgtacggege etetgacage aacgtetatg acetectaaa ggacetagag gaaggeatee aaacgetgat ggggaggetg gaagatggea geeeeeggae tgggeagate tteaageaga cetacageaa gttegacaca aacteacaca acgatgacg actactcaag aactacggge tgetetactg etteaggaag gacatggaca aggtegagae atteetgege 54</pre> | <211> <212> | 588 DNA | | | | | |
| catatgttcc caaccattce cttatecagg ctttttgaca acgetatget cegegeceat cgtctgcacc agetggeett tgacacctae caggagtttg aagaageeta tateccaaag 12 gaacagaagt atteatteet geagaaceee cagaceteee tetgtttete agagtetatt 18 ccgacaccet ccaacaggga ggaaacacaa cagaaateea acetagaget geteegeate 24 teeetgetge teatecagte gtggetggag eeegtgeagt teeteaggag tgtettegee 30 aacageetgg tgtacggege etetgacage aacgtetatg aceteetaaa ggacetagag 36 gaaggeatee aaacgetgat ggggaggetg gaagatggea geeeeeggae tgggeagate 42 tteaageaga cetacageaa gttegacaca aacteacaca acgatgacg actacteaag 48 aactacggge tgetetactg etteaggaag gacatggaca aggtegagae atteetgge 54 aactacggge tgetetactg etteaggaag gacatggaca aggtegagae atteetgege 54 | | synthetic seque | ence | | . 3 | | |
| gaacagaagt attcattcct gcagaacccc cagacctccc tetgtttctc agagtctatt 18 ccgacaccct ccaacaggga ggaaacacaa cagaaatcca acctagagct gctccgcatc 24 tccctgctgc tcatccagtc gtggctggag cccgtgcagt tcctcaggag tgtcttcgcc 36 aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag 36 gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 42 ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacg actactcaag 48 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 | | • | cttatccagg | ctttttgaca | acgctatgct | ccgcgcccat | 60 |
| ccgacaccct ccaacagga ggaaacacaa cagaaatcca acctagagct gctccgcatc tccctgctgc tcatccagtc gtggctggag cccgtgcagt tcctcaggag tgtcttcgcc 36 aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag gaaggcatcc aaacgctgat ggggaggctg gaagatggca gccccggac tgggcagatc ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 | cgtctg | cacc agetggeett | tgacacctac | caggagtttg | aagaagccta | tatcccaaag | 120 |
| tecetgetge teatecagte gtggetggag eeegtgeagt teeteaggag tgtettegee 30 aacageetgg tgtacggege etetgacage aacgtetatg acetectaaa ggacetagag 36 gaaggeatee aaacgetgat ggggaggetg gaagatggea geeeeeggae tgggeagate 42 tteaageaga eetacageaa gttegacaca aacteacaca acgatgaeg actaeteaag 48 aactaeggge tgetetactg etteaggaag gacatggaea aggtegagae atteetgege 54 | gaacag | aagt attcattcct | gcagaacccc | cagacetece | tctgtttctc | agagtctatt | 180 |
| aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag 36 gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 42 ttcaagcaga cctacagcaa gttcgacaca aactcacada acgatgacgc actactcaag 48 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 | ccgaca | ccct ccaacaggga | ggaaacacaa | cagaaatcca | acctagagct | gctccgcatc | 240 |
| gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 42 ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 48 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 | tacatg | ctgc tcatccagtc | gtggctggag | cccgtgcagt | tcctcaggag | tgtcttcgcc | 300 |
| ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 | aacagc | ctgg tgtacggcgc | ctctgacagc | aacgtctatg | acctcctaaa | ggacctagag | 360 |
| aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 54 | gaaggc | atcc aaacgctgat | ggggaggctg | gaagatggca | gcccccggac | tgggcagatc | 420 |
| | ttcaag | caga cctacagcaa | gttcgacaca | aactcacaca | acgatgacgc | actactcaag | 480 |
| atcgtgcagt gccgctctgt ggagggcagc tgtggcttct agggatcc 58 | aactac | gggc tgctctactg | cttcaggaag | gacatggaca | aggtcgagac | attcctgcgc | 540 |
| | atcgtg | cagt gccgctctgt | ggagggcagc | tgtggcttct | agggatcc | | 588 |

<210> 8 <211> 192

```
PRT
<212>
      Artificial
<213>
<220>
<223>
     synthetic sequence
<220>
<221>
      mat peptide
<222>
      (1)..()
<400>
Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu
Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
                               25
Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn
Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn
Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser
            Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser
Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
                               105
Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg
                           120
Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr
                       135
Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn
                                       155
                   150
Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr
Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe
       9
<210>
<211>
       2907
<212>
       DNA
<213>
       Artificial
<220>
       synthetic sequence
<223>
```

<400> 9

| ccggatațag | tteeteettt | cagcaaaaaa | cccctcaaga | cccgtttaga | ggccccaagg | 60 | |
|------------|------------|------------|------------|------------|--------------|------|---|
| ggttatgcta | gttattgctc | agcggtggca | gcagccaact | cagcttcctt | tcgggctttg | 120 | |
| tttagcagcc | taggtattaa | tcaattagtg | gtggtggtgg | tggtggtggt | gctcgagtgc | 180 | |
| ggccgcaagc | ttgtcgacgg | agctcgcctg | caggcgcgcc | aaggcctgta | cagaattcgg | 240 | |
| atccccgata | tcccatggga | ctcttgtcgt | cgtcatcacc | ggagccacca | ccggtaccca | 300 | - |
| gatctgggct | gtccatgtgc | tggcgttcga | atttagcagc | agcggtttct | ttcataccaa | 360 | |
| ttgcagtact | accgcgtggc | accagacccg | cggagtgatg | gtgatggtga | tgaccagaac | 420 | |
| cactagtaca | cacatatgta | tatctccttc | ttaaagttaa | acaaaattat | ttctagaggg | 480 | |
| gaattgttat | ccgctcacaa | ttcccctata | gtgagtcgta | ttaatttcgc | gggatcgaga | 540 | |
| tegatetega | tcctctacgc | cggacgcatc | gtggccggca | tcaccggcgc | cacaggtgcg | 600 | |
| gttgctggcg | cctatatcgc | cgacatcacc | gatggggaag | atcgggctcg | ccacttcggg | 660 | |
| ctcatgagcg | cttgtttcgg | cgtgggtatg | gtggcaggcc | ccgtggccgg | gggactgttg | 720 | |
| ggcgccatct | ccttgcatgc | atggcgtaat | catggtcata | gctgtttcct | gtgtgaaatt | 780 | |
| gttatccgct | cacaattcca | cacaacatac | gagccggaag | cataaagtgt | aaagcctggg | 840 | |
| gtgcctaatg | agtgagctaa | ctcacattaa | ttgcgttgcg | ctcactgccc | gctttccagt | 900 | - |
| cgggaaacct | gtcgtgccag | ctgcattaat | gaatcggcca | acgcgcgggg | agaggcggtt ' | 960 | |
| tgcgtattġg | gcgctcttcc | gcttcctcgc | tcactgactc | gctgcgctcg | gtcgttcggc | 1020 | 7 |
| tgcggcgagc | ggtatcagct | cactcaaagg | cggtaatacg | gttatccaca | gaatcagggg | 1080 | |
| ataacgcagg | aaagaacatg | tgagcaaaag | gccagcaaaa | ggccaggaac | cgtaaaaagg | 1140 | |
| ccgcgttgct | ggcgtttttc | cataggctcc | gcccccctga | cgagcatcac | aaaaatcgac | 1200 | |
| gctcaagtca | gaggtggcga | aacccgacag | gactataaag | ataccaggcg | tttccccctg | 1260 | |
| gaagctccct | cgtgcgctct | cctgttccga | ccctgccgct | taccggatac | ctgtccgcct | 1320 | |
| ttctcccttc | gggaagcgtg | gcgctttctc | atagçtcacg | ctgtaggtat | ctcagttcgg | 1380 | |
| tgtaggtcgt | tegetecaag | ctgggctgtg | tgcacgaacc | ccccgttcag | cccgaccgct, | 1440 | |
| gcgccttatc | cggtaactat | cgtcttgagt | ccaacccggt | aagacacgac | ttatcgccac | 1500 | |
| tggcagcagc | cactggtaac | aggattagca | gagcgaggta | tgtaggcggt | gctacagagt | 1560 | |
| tcttgaagtg | gtggcctaac | tacggctaca | ctagaaggac | agtatttggt | atctgcgctc | 1620 | |
| tgctgaagcc | agttaccttc | ggaaaaagag | ttggtagctc | ttgatccggc | aaacaaacca | 1680 | |
| ccgctggtag | cggtggtttt | tttgtttgca | agcagcagat | tacgcgcaga | aaaaaaggat | 1740 | |

```
ctcaagaaga tootttgato ttttctacgg ggtctgacgc tcagtggaac gaaaactcac
                                                                     1800
gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt
                                                                     1860
aaaaatgaag tittaaatca atctaaagta tatatgagta aacttggtct gacagttacc
                                                                     1920
aatgettaat cagtgaggea cetateteag egatetgtet atttegttea tecatagttg
                                                                     1980
cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg
                                                                     2040
ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc
                                                                     2100
cagcoggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta
                                                                     2160
ttaattgttg ccgggaaget agagtaagta gttcgccagt taatagtttg cgcaacgttg
                                                                     2220
ttgecattge tacaggeate gtggtgtcae getegtegtt tggtatgget teatteaget
                                                                     2280
ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta
                                                                     2340
gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg
                                                                     2400
ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga
                                                                     2460
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt
                                                                     2520
                                                                     2580
gcccggcgtc aatacgggat aataccgcgc cacatagcag aactitaaaa gtgctcatca
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt
                                                                     2640
cgatgtaacc cactogtgca cocaactgat ottoagoato tittactito accagogtit
                                                                     2700
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga
                                                                     2760
                                                                     2820
aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat cagggttatt
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc
                                                                     2880
gcacatttcc ccgaaaagtg ccacctg
                                                                     2907
```

```
<210> 10
```

cgccatatga aaaagacagc tatcgcgatt gcagtggcac tggctggttt cgctaccgta

60

<211> 73

<212> DNA

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> sig_peptide

<222> (7)..(69)

<400> 10

```
<210>
      11
<211>
      21
<212> PRT
<213> Escherichia coli
<220>
<221> SIGNAL
<222> (1)..(21)
<400> 11
Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
Thr Val Ala Gln Ala
<210> 12
<211> 42
<212> DNA
<213> Artificial
<220>
<223> synthetic sequence
ggacatatgc tgaagctttc ccaaccattc ccttatccag gc
                                                                     42
<210> 13
<211> 28
<212> DNA
<213> Artificial
<220>
<223> synthetic sequence
<400> 13
cgcggatccc tccacagagc ggcactgc
                                                                     28
<210> 14
<211> 578
<212> DNA
<213> Artificial
<220>
<223> synthetic sequence
<400> 14
catatgctga agctttccca accattccct tatccaggct ttttgacaac gctatgctcc
gcgcccatcg tctgcaccag ctggcctttg acacctacca ggagtttgaa gaagcctata
                                                                   120
```

73

gcgcaagctt gag

| tcçcaaagga | acagaagtat | tcattcctgc | agaaccccca | gacctccctc | tgtttctcag | 180 |
|------------|------------|------------|------------|-------------|------------|-----|
| agtctattcc | gacaccctcc | aacagggagg | aaacacaaca | gaaatccaac. | ctagagctgc | 240 |
| tccgcatctc | cctgctgctc | atccagtcgt | ggctggagcc | cgtgcagttc | ctcaggagtg | 300 |
| tcttcgccaa | cagcctggtg | tacggcgcct | ctgacagcaa | cgtctatgac | ctcctaaagg | 360 |
| acctagagga | aggcatccaa | acgctgatgg | ggaggctgga | agatggcagç | ccccggactg | 420 |
| ggcagatctt | caagçagacc | tacagcaagt | tcgacacaaa | ctcaçacaac | gatgacgcac | 480 |
| tactcaagaa | ctacgggctg | ctctactgct | tcaggaagga | catggacaag | gtcgagacat | 540 |
| tcctgcgcat | cgtgcagtgc | cgctctgtgg | agggatcc | | | 578 |
| | | | | | | |

<211> 630

<212> DNA

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> source

<222> (4)..(66)

<223> Eschericia coli

<220>

<221> source

<222> (67)..(630)

<223> homo sapiens

<400> 15

catatgaaaa agacagctat cgcgattgca gtggcactgg ctggtttcgc taccgtagcg caagetttee caaceattee ettatecagg etttttgaca acgetatget eegegeecat 120 cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag 180 gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt 240 ccgacaccet ccaacaggga ggaaacacaa cagaaatcca acctagaget gctccgcatc 300 tecetgetge teatecagte gtggetggag ecegtgeagt teeteaggag .tgtettegee 360 aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag 420 gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 480 ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 540 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 600

```
<210> 16 <211> 208
```

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> SIGNAL

<222> (1)..(21)

<223> from Eschericia coli

<400> 16

Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala 1 5 10 15

Thr Val Ala Gln Ala Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp 20 25 30

Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr 35 40 45

Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser 50 60

Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro 65 70 75 80

Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu 85 90 95

Leu Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln
100 105 110

Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp 115 120 125

Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr 130 135 140

Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe 145 150 155 160

Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala 165 170 175

Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp 180 185

Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly

300.

| <210><211><212><212><213> | 17 570 DNA Artif | ficia | ıl | | | | | | | | | | | |
|---------------------------|---------------------------|-----------|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| <220> <223> | syntl | netic | seq | uenc | e | | | | | | | | | |
| <400> | 17 | | | | | | | | | | | | | |
| catatg | ttcc (| caaco | atto | c ct | tato | cagg | ctt | tttg | jaca | acgo | tatç | gct | ccgcç | cccat |
| cgtctg | cacc | agcto | gcct | t te | Jacac | ctac | cag | gagt | ttg | aaga | agco | ta | tatco | caaag |
| gaacag | aagt a | attca | ttcc | t go | agaa | cccc | cag | jacct | ccc | tcts | tttc | ctc | agagt | ctatt |
| ccgaca | ccct (| ccaac | caggg | a gg | jaaac | cacaa | cag | jaaat | cca | acct | agag | gct | gctc | gcatc |
| tccctg | ctgc t | cato | ccagt | c gt | ggct | ggag | ccc | gtgc | agt | tcct | cago | gag | tgtct | tcgcc |
| aacagc | ctgg I | tgtac | ggcg | c ct | ctga | acago | aac | gtct | atg | acct | ccta | aaa | ggaco | tagag |
| gaaggc | atcc a | aaacg | gctga | t gg | ggag | gctg | gaa | gato | ggca | gcc | ccgg | jąc | tgggd | agatc |
| ttcaag | cąga (| cctac | cagca | a gt | tcga | acaca | aac | tcac | caca | acga | tgad | gc | actac | tcaag |
| aactac | gggc 1 | tgcto | ctact | g ct | tcag | ggaac | gad | atgo | gaça | aggt | cgaç | gac | attco | tgcgc |
| atcgtg | cagt (| gccgo | etete | ıt go | jaggg | gated | | | | , | | | | |
| | ; | | | | | | | | | | | - | | |
| <210> <211> | 18 188 | | | | | | | | | | | | | • |
| <212> | PRT Arti | ficia | al | | | | | | | | | ٠. | | |
| | | | | | | • | | | | | | | | |
| <220> <223> | synt | hetio | c sec | ueno | ce | | | | | | - | | | |
| <400> | 18 | | | | | | | | | | | ÷ | | |
| | | _ | _ | | | | | | | | | | | |
| Met Ph 1 | e Pro | Thr | Ile 5 | Pro | Leu | Ser | Arg | Leu 10 | Phe | Asp | Asņ | Ala | Met 15 | Leu |
| Arg Al | a His | Arg 20 | Leu | His | Gln | Leu | Ala 25 | Phe | Asp | Thr | Tyr | Gln 30 | | Phe |
| Glu Gl | u Ala 35 | Tyr | Ile | Pro | Lys | Glu 40 | Gln | Lys | Tyr | Ser | Phe 45 | Leu | Gln | Asn |
| Pro Gl 50 | | Ser | Leu | Cys | Phe 55 | Ser | Glu | Ser | Ile | Pro 60 | Thr | Pro | Ser | Asn |
| Arg Gl | u Glu | Thr | Gln | Gln 70 | Lys | Ser | Asn | Leu | Glu 75 | Leu | Leu | Arg | Ile | Ser 80 |

| Leu ' | Leu | Leu | Ile | Gln 85 | Ser | Trp | Leu | Glu | Pro 90 | Val | Gln | Phe | Leu · | Arg 95 | Ser | | | |
|--|--|--|---|---|--|--|---|---|--|--------------------------------------|-------------------------------------|--|--|--------------------------------------|----------------------------------|----------------------|----------------------------|----------------------|
| Val | Phe | Ala | Asn 100 | Ser | Leu | Val | Tyr | Gly 105 | Ala | Ser | Asp | Ser | Asn 110 | Val | Tyr | | | |
| Asp | Leu | Leu 115 | Lys | Asp | Leu | Glu | Glu 120 | Gly | Ile | Gln | Thr | Leu 125 | Met | Gly | Arg | | | |
| Leu | Glu 130 | Asp | Gly | Ser | Pro | Arg 135 | Thr | Gly | Gln | Ile | Phe 140 | Lys | Gln | Thr | Tyr | | | |
| Ser 145 | Lys | Phe | Asp | Thr | Asn 150 | Ser | His | Asn | Asp | Asp 155 | Ala | Leu | Leu | Lys | Asn 160 | | | |
| Tyr | Gly | Leu | Leu | Tyr 165 | Cys | Phe | Arg | Lys | Asp 170 | Met | Asp | Lys | Val | Glu 175 | Thr | | • | |
| Phe | Leu · | Arg | Ile 180 | Val | Gln | Cys | Arg | Ser 185 | Val | Glu | Gly | | | | | | | , |
| <210 <211 | | 19 52 | ٠ | | | | | | | | | • | | | | | | |
| <212 <213 | | ONA Arti: | ficia | al | | | | | | | • | | | | | | | |
| <220 <223 | | syntl | netio | c sec | quenc | ce | • | ÷ | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | |
| <400 | | 19 Fair 4 | aata: | à tho à t | ta ta | racti | tett. | d da | 2200 | atta | ccti | t a tro | חבר | ac | • | | • | 52 |
| | | | catga | atcal | tg tạ | ggct | tetto | d dda | aacc | attc | ccti | tatc | cag | gc | •. | | | 52 |
| <210 | catai | tga (20 | catga | ātcai | tg tạ | gget | tetto | d dda | aacc | attc | ccti | tatc | cag | gc | | | | 52 |
| taco |)> : | tga (| catga | ātcai | tg to | gget | tetto | d dda | aacc | attc | ccti | tatc | cag | gc | | - | | 52 |
| <210 <211 |)> 1 l> ! 2> I | tga (20 588 | | | tg tọ | gget | tetto | d cca | aacc | attc | ccti | tatc | cag | gc | | | | 52 |
| <210 <211 <212 | catai)> : 1> ! 2> I 3> ! | tga (20 588 DNA Arti: | ficia | al | | | tette | d dda | aacc | attc | ccti | tato | cag | gc | | | | 52 |
| <210 <211 <212 <213 | catai)> : 1> ! 2> I 3> ! | tga (20 588 DNA Arti: | ficia | | | | tette | e ee | aacc | attc | ccti | tatco | cag | gc | | | | 52 |
| <210 <211 <212 <213 <223 <400 | Catal O> 2 L> ! 2> I O> 3 >> 3 >> 3 >> 3 | tga (20 588 DNA Arti: synt) | ficia hetio | al c sec | quenc | ce | | <u>,</u> 3 | | | | | | | gacaa | ac | | 52 |
| <210 <211 <212 <213 <223 <400 cata | 0> 2 L> ! 2> I 3> 2 0> 3 | tga (20 588 DNA Arti syntl | ficia hetio gatea | al c sec | gg ct | ce | tccca | a ace | catto | ccct | tate | ccag | gct | ttt | gacaa | | | |
| <210 <211 <212 <213 <223 <400 cata | 0> 2 1> ! 2> I 3> 2 0> 3 0> 3 | 20 588 DNA Arti: syntl | ficia hetio gatoa gegeo | al c sec atgto | quenc gg ct | ce ctct | tccca | a ace | catte | cect | tate | ccag | gct cca | tttt | | aa | 12 | 60 |
| <210 <211 <212 <223 <223 <400 cata gcta | Catal O> 2 L> ! O> 3 O> 3 Atgac atgct | tga (20 588 DNA Arti: syntl 20 cat (tcc (| ficia hetic gatca gegeo tecca | al c sec atgto ccato aaago | gg ct cg to ga ao | ce ctgc caga | tccca acca agtal | a acc | catte ggcc | ccct tttg ctgc | tate acae agae | ccag ccta accc | gct cca cca | tttt ggag gacc | tttga | aa tc | 12 | 60 20 |
| <210 <211 <212 <223 <400 cata gcta gaag tgtt | catal O> | 20 588 DNA Arti: syntl 20 cat (| ficia netion gatea gegeo tecca agte | al c sec atgto ccato aaagg | quenc gg ct cg to ga ac cc ga | ce ctgc caga | tccca acca agta cctco | a ace g cte t tca c aae | catte ggcc attc cagg | ccct tttg ctgc gagg | tato acao agao aaao | ccag ccta accc acca | gct cca cca | tttt ggag gacc gaaa | tttg: | aa tc ac | 12 | 60 20 80 |
| <210 <211 <212 <223 <400 cata gcta gaag tgtt | catal O> 2 L> ! O> 3 O> 3 O> 3 atgain atgain ctctc | tga (20 588 DNA Arti: synt) cat (ata cag (| ficia netion gatea gegeo tecea agtea teego | al c sec atgto ccato aaagg tatto | queno gg ct cg to ga ao cc ga tc co | ce ctgc caga acaco | tccca acca agta cctcc tgct | a aco g cto t tco c aao c ato | catto ggcc attc caggo ccag | ccct tttg ctgc gagg tcgt | tato acao agao aaao ggc | ccago cctao accco cacao tgga | gct cca cca aca gcc | tttt ggag gacc gaaa cgtg | tttga tccci | aa tc ac | 111 11 24 30 | 60 20 80 |
| <210 <211 <212 <221 <222 <400 cata gcta gaag tgtt ctag ctca | catal O> 2 L> ! L> ! O> 3 S | tga (20 588 DNA Arti: syntl 20 cat (tcc (ata cag (tgc) | ficia hetio gatea gegeo tecea agtea tecego tecto | al c sec atgto ccato aaagg tatto cato | queno gg ch cg to ga ao co ga to co | ce ctgca cagaa acaca ctgca | tecea accaq agtal cetec tgctc | a aco g cto t tco c aao c ato | catto ggcc atto caggo ccag | ccct tttg ctgc gagg tcgt | tate acae agae aaae ggc | ccag ccta accc caca tgga acag | gct cca cca aca gcc caa | tttt ggag gacc gaaa cgtg | tttga tccci tccaa cagti | aa tc ac tc | 12 14 24 30 36 | 50 20 80 40 |

<211> 191

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<400> 21

Ser Cys Gly Phe Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn 1 5 10 15

Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr
20 25 30

Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe 35 40 45

Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr 50 55 60

Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu 65 70 75 80

Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe
85 90 95

Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser 100 105 110

Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu 115 120 125

Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys 130 135 140

Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu 145 150 155 160

Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys
165 170 175

Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly
180 185 190

<210> 22

<211> 42

<212> DNA

<213> Artificial

```
<220>
<223> synthetic sequence
<400> 22
tacgaattcc attgatcatg tggcttctag taggtcgacg at
                                                                    42
<210> 23
<211>
<212> PRT
<213> Artificial
<220>
<223> synthetic sequence
<400> 23
Ser Cys Gly Phe
<210> 24
<211> 51
<212> DNA
<213> Artificial
<220>
<223> synthetic sequence
<400> 24
                                                                    51
tacgaattcc attgatcatg tggcttcaaa aagaaatagt aggtcgacga t
<210> 25
<211> 7
<212> PRT
<213> Artificial
<220>
<223> synthetic sequence
<400> 25
Ser Cys Gly Phe Lys Lys
<210> 26
<211> 1161
<212> DNA
<213> Artificial
<220>
<223> synthetic sequence
<400> 26
catatgacat gatcatgtgg cttcttccca accattccct tatccaggct ttttgacaac
                                                                    60
gctatgctcc gcgcccatcg tctgcaccag ctggcctttg acacctacca ggagtttgaa
                                                                   120
```

| gaagcctata | tcccaaagga - | acagaagtat | tcattcctgc | agaaccccca | gacctccctc | 180 |
|------------|-----------------|------------|------------|------------|------------|------|
| tgtttctcag | agtctattcc | gacaccctcc | aacagggagg | aaacacaaca | gaaatccaac | 240 |
| ctagagctgc | tccgcatctc | cctgctgctc | atccagtcgt | ggctggagcc | cgtgcagttc | 300 |
| ctcaggagtg | tcttcgccaa | cagcctggtg | tacggcgcct | ctgacagcaa | cgtctatgac | 360 |
| ctcctaaagg | acctagagga | aggcatccaa | acgctgatgg | ggaggctgga | agatggcagc | 420 |
| ccccggactg | ggcagatctt | caagcagacc | tacagcaagt | tcgacacaaa | ctcacacaac | 480 |
| gatgacgcac | tactcaagaa | ctacgggctg | ctctactgct | tcaggaagga | catggacaag | 540 |
| gtcgagacat | tcctgcgcat | cgtgcagtgc | cgctctgtgg | agggatcatg | tggcttcttc | 600 |
| ccaaccattc | ccttatccag | gctttttgac | aacgctatgc | tccgcgccca | tcgtctgcac | 660 |
| cagetggeet | ttgacaccta | ccaggagttt | gaagaagcct | atatcccaaa | ggaacagaag | 720 |
| tattcattcc | tgcagaaccc | ccagacctcc | ctctgtttct | cagagtctat | teegacacee | 780 |
| tccaacaggg | aggaaacaca | acagaaatcc | aacctagagc | tgctccgcat | ctccctgctg | 840 |
| ctcatccagt | cgtggctgga | gcccgtgcaġ | ttcctcagga | gtgtcttcgc | caacagcctg | 9Ó0 |
| gtgtacggcg | cctctgacag | caacgtctat | gacctcctaa | aggacctaga | ggaaggcatc | 960 |
| caaacgctga | tggggaggct | ggaagatggc | agcccccgga | ctgggcagat | cttcaagcag | 1020 |
| acctacágca | agttcgacac | aaactcacac | aacgatgacg | cactactcaa | gaactacggg | 1080 |
| ctgctctact | gcttcaggaa | ggacatggac | aaggtcgaga | cattcctgcg | catcgtgcag | 1140 |
| tgccgctctg | tggagggatc | C | | | | 1161 |

<211> 382

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<400> 27

Ser Cys Gly Phe Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn 1 5 10 15

Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr 20 25 30

Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe 35 40 45

| Leu | Gln 50 | Asn | Pro | Gln | Thr | Ser 55 | Leu | Cys | Phe | Ser | Glu 60 | Sẹr | Ile | Pro | Thr |
|------------|------------|------------|-------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------|
| Pro 65 | Ser | Asn | Arg | Glu | Glu 70 | Thr | Gln | Gln | Lys | Ser 75 | Asn | Leu | Glu | Leu | Leu 80 |
| Arg | Ile | Ser | Leu | Leu 85 | Leu | Ile | Gln | Ser | Trp 90 | Leu | Ģlu | Pro | Val | Gln 95 | Phe |
| Leu | Arg | Ser | Val [°] 100 | Phe | Ala | Asn | Ser | Leu 105 | Val | Tyr | Gly | Ala | Ser 110 | Asp | Ser |
| Asn | Val | Tyr 115 | Asp | Leu | Leu | Lys | Asp 120 | Leu | Glu | Glu | Gly | Ile 125 | Gln | Thr | Leu |
| Met | Gly 130 | Arg | Leu | Glu | Asp | Gly 135 | Ser | Pro | Arg | Thr | Gly 140 | Gln | Ile | Phe | Lys |
| Gln 145 | Thr | Tyr | Ser | Lys | Phe 150 | Asp | Thr | Asn | Ser | His 155 | | Asp | Asp | Ala | Leu . 160. |
| Leu | Lys | Asn | Tyr | Gly 165 | Leu | Leu | Tyr | | Phe 170 | Arg | Lys | Asp | Met | Asp. | Lys |
| Val | Gľu | Thr | Phe 180 | Leu | Arg | Ile | Val | Gln 185 | Cys | Arg | Ser | Val | Glu 190 | Gly | .ser |
| Cys | Gly | Phe 195 | Phe | Pro | Thr | Ile | Pro 200 | Leu | Ser | Arg | Leu | Phe 205 | Asp | Asn | Ala |
| Met | Leu 210 | Arg | Alá | His | Arg | Leu 215 | His | Gln | Leu | Ala | Phe 220 | Asp | Thr | Tyr | Gln |
| Glu 225 | Phe | Glu | Glu | Ala | Tyr 230 | Ile | Pro | Lys | Glu | Gln 235 | Lys | Tyr | Ser | Phe | Leu 240 |
| Gln | Asn | Pro | Glń | Thr 245 | Ser | Leu | Cys | Phe | Ser 250 | Glu | Ser | Ile | Pro | Thr 255 | Pro |
| | Asn | Arg | Glu 260 | Glu | Thr | GÌn | Gln | Lys 265 | Ser | Asn | Leu | Glu | Leu 270 | Leu | Arg |
| Ile | Ser | Leu 275 | | Leu | Ile | Gln | Ser 280 | | Leu | Glu | Pro | Val 285 | | Phe | Leu |
| Arg | Ser 290 | | Phe | Ala | Asn | Ser 295 | | Val | Tyr | | Ala 300 | Ser | Asp | Ser | Asn |
| Val 305 | _ | Asp | Ļeu | Leu | Lys 310 | | Leu | Glu | Glu | Gly 315 | | Gln | Thr | Leu | Met 320 |
| Gly | Arg | Leu | Glu | Asp 325 | | Ser | Pro | Arg | Thr 330 | | Gln | Ile | Phe | Lys 335 | Gln |
| Thr | Tyr | Ser | Lуs 340 | | Asp | Thr | Asn | Ser 345 | | Asn | Asp | Asp | Ala 350 | | Leu |

Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val 355 360 365

Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly 370 375 380

<210> 28

<211> 1152

<212> DNA

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> misc feature

<222> (574)..(1146)

<223> sequence is repeated N-1 times, where N is a positive whole numbe

<400> tgatcatgtg gettetteec aaceatteec ttatecagge tttttgacaa egetatgete 60 egegeceate gtetgeacea getggeettt gacacetace aggagtttga agaageetat 120 atcccaaagg aacagaagta ttcattcctg cagaaccccc agacctccct ctgtttctca 180 gagtetatte egacaceete caacagggag gaaacacaac agaaatecaa eetagagetg 240 ctccgcatct ccctgctgct catccagtcg tggctggagc ccgtgcagtt cctcaggagt 300 gtettegeca acageetggt gtaeggegee tetgaeagea aegtetatga eeteetaaag 360 420 gacctagagg aaggcatcca aacgctgatg gggaggctgg aagatggcag cccccggact 480 gggcagatet teaageagae etacageaag ttegacacaa aeteacacaa egatgaegea ctactcaaga actacgggct gctctactgc ttcaggaagg acatggacaa ggtcgagaca 540 ttcctgcgca tcgtgcagtg ccgctctgtg gagggatcat gtggcttctt cccaaccatt 600 cccttatcca ggctttttga caacgctatg ctccgcgccc atcgtctgca ccagctggcc 660 tttgacacct accaggagtt tgaagaagcc tatatcccaa aggaacagaa gtattcattc 720 ctgcagaacc cccagacctc cctctgtttc tcagagtcta ttccgacacc ctccaacagg 780 840 gaggaaacac aacagaaatc caacctagag ctgctccgca tctccctgct gctcatccag tegtggetgg agecegtgea gtteeteagg agtgtetteg ceaacageet ggtgtaegge 900 geetetgaca geaacgteta tgacetecta aaggacetag aggaaggeat ceaaacgetg 960 1020 atggggagge tggaagatgg cageeeeegg actgggeaga tetteaagea gaeetacage 1080 aagttegaca caaacteaca caacgatgac gcactactea agaactacgg getgetetac

tgcttcagga aggacatgga caaggtcgag acattcctgc gcatcgtgca gtgccgctct 1140
gtggagggat cc 1152

<210> 29

<211> 382

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> MISC_FEATURE

<222> (191)..(381)

<223> sequence is repeated N-1 times, where N is a positive whole numbe

<400> 29

Ser Cys Gly Phe Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn 1 5 10 : 15

Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr 20 25 30

Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe 35 40 45

Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr 50 55 60

Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu 65 70 75 80

Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe 85 90 95

Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser
100 105 110

Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu 115 120 125

Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys 130 135 140

Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys
165 170 175

Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser

180 185 190

| Cys | Gly | Phe 195 | Phe | Pro | Thr | Ile | Pro 200 | Leu | Ser | Arg | Leu | Phe 205 | Asp | Asn | Ala | |
|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|
| Met | Leu 210 | Arg | Ala | His | Arg | Leu 215 | His | Ģln | Leu | Ala | Phe 220 | Asp | Thr | Tyr | Glņ | i. |
| Glu 225 | Phe | Glu | Glu | Ala | Tyr 230 | Ile | Pro | Lys | Glu | Gln 235 | Lys | Tyr | Ser | Phe | Leu 240 | |
| Gln | Asn | Pro | Gln | Thr 245 | Ser | Leu | Cys | Phe | Ser 250 | Glu | Ser | Ile | Pro | Thr 255 | Pro | |
| Ser | Asn | Arg | Glu 260 | Glu | Thr | Gln | Gln | Lys 265 | Ser | Asn | Leu | Glu | Leu 270 | Leu | Arg | |
| Ile | Ser | Leu 275 | Leu | Leu | Ile | Gln | Ser 280 | Trp | Leu | Glu | Pro | Val 285 | Gln | Phe | Leu | |
| Arg | Ser 290 | Val | Phe | Ala | Asn | Ser 295 | Leu | Val | Tyr | Gly | Ala 300 | Ser | Asp | Ser | Asn | |
| Val 305 | Tyr | Asp | Leu | Leu | Lys 310 | Asp | Leu | Glu | .Glu | Gly 315 | Ile | Gln | Thr | Leu | Met 320 | |
| Gly | Arg | Leu | Glu | Asp 325 | Gly | Ser | Pro | Arg | Thr 330 | Gly | Gln | Ile | Phe | Lys 335 | Gln | |
| Thr | Tyr | Ser | Lys 340 | Phe | Asp | Thr | Asn | Ser 345 | His | Asn | Asp | Asp | Ala 350 | Leu | Leu | |
| Lys | Asn | Tyr 355 | Gly | Leu | Leu | Tyr | Cys 360 | Phe | Arg | Lys | Asp | Met 365 | Asp | Lys | Val | |
| Glu | Thr 370 | Phe | Leu | Arg | Ile | Val 375 | Gln | Cys | Arg | Ser | Val 380 | Glu | Gly | | | |
| <21 <21 | | 30 606 | | | | | | | | | | | | | | |
| <21 | 2>] | DNA | <u> </u> | 7 | | | | | | | | | | | | |
| .<21 | 3> 1 | Artı | fici | a⊥ | | | ٠ | | | | | | | | | |
| <22 <22 | | synt] | heti | c se | quen | ce | | | | | | | | | | |
| <40 cat | | 30 tcc (| caac | catt | cc c | ttat | ccag | g cti | tttt | gaca | acg | ctat | gct | ccgc | gcccat | 60 |
| cgt | ctgc | acc a | agct | ggcc | tt t | gaca | ccta | c ca | ggag | tttg | aag | aagc | cta | tatc | ccaaag | 120 |
| gaạ | caga | agt (| attc | attc | ct g | caga | accc | c ca | gacc | tccc | tct | gttt | ctc | agag | tctatt | 180 |
| ccg | acaç | cct (| ccaa | cagg | ga g | gaaa | caca | a ca | gaaa | tcca | acc | taga | gct | gctc | cgcatc | 240 |
| tcc | ctgc | tgc | tcat | ccag | tc g | tggc | tgga | g cc | cgtg | cagt | tcc | tcag | gag | tgtc | ttcgcc | 300 |

aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag 360
gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 420
ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 480
aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 540
atcgtgcagt gccgctctgt ggagggatcc gaattccatt gatcatgtgg cttctagtag 600
gtcgac 606

<210> 31

<211> 1737

<212> DNA

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> misc_feature

<222> (1138)..(1710)

<223> sequence is repeated N-1 times, where N is a positive whole numbe

<400> 31 catatgttcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgcccat 60 120 cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag 180 gaacagaagt atteatteet geagaaceee cagaceteee tetgtttete agagtetatt 240 cegacacect ccaacaggga ggaaacacaa cagaaatcca acctagaget geteegcate 300 tecetgetge teatecagte gtggetggag ecegtgeagt teeteaggag tgtettegee aacageetgg tgtaeggege etetgaeage aacgtetatg aceteetaaa ggaeetagag 360 gaaggcatce aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 420 ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 480 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 540 atcgtgcagt gccgctctgt ggagggatca tgtggcttct tcccaaccat tcccttatcc 600 aggetttttg acaacgetat geteegegee categtetge accagetgge etttgacace 660 taccaggagt ttgaagaagc ctatatccca aaggaacaga agtattcatt cctgcagaac 720 ccccagacct ccctctgttt ctcagagtct attccgacac cctccaacag ggaggaaaca 780 caacagaaat ccaacctaga getgeteege atetecetge tgeteateca gtegtggetg 840

```
gagecegtge agttecteag gagtgtette gecaacagee tggtgtaegg egeetetgae
                                                                     900
agcaacgtct atgacctcct aaaggaccta gaggaaggca tccaaacgct gatggggagg
                                                                     960
ctggaagatg gcagccccg gactgggcag atcttcaagc agacctacag caagttcgac
                                                                     1020
acaaactcac acaacgatga cgcactactc aagaactacg ggctgctcta ctgcttcagg
                                                                     1080
aaggacatgg acaaggtcga gacattcctg cgcatcgtgc agtgccgctc tgtggaggga
                                                                     1140
tcatgtggct tcttcccaac cattccctta tccaggcttt ttgacaacgc tatgctccgc
                                                                     1200
geceategte tgeaceaget ggeetttgae acetaceagg agtttgaaga ageetatate
                                                                     1260
ccaaaggaac agaagtattc attcctgcag aacccccaga cctccctctg tttctcagag
                                                                     1320
totattooga caccotocaa cagggaggaa acacaacaga aatocaacot agagetgoto
                                                                     1380
cgcatctccc tgctgctcat ccagtcgtgg ctggagcccg tgcagttcct caggagtgtc
                                                                     1440
ttegecaaca geetggtqta eggegeetet gacageaacg tetatgaeet eetaaaggae
                                                                     1500
ctagaggaag gcatccaaac gctgatgggg aggctggaag atggcagccc ccggactggg
                                                                     1560
cagatettea ageagaceta cageaagtte gacacaaact cacacaacga tgacgcacta
                                                                     1620
ctcaagaact:acgggctgct ctactgcttc aggaaggaca tggacaaggt cgagacattc
                                                                     1680
ctgcgcatcg tgcagtgccg ctctgtggag ggatcatgtg gcttctagta ggtcgac
                                                                     1737
```

```
<210> 32
<211> 574
<212> PRT
```

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> MISC_FEATURE

<222> (379)..(569)

<223> sequence is repeated N-1 times, where N is a positive whole numbe

<220>

<221> mat_peptide

<222> (1)..()

<400> 32

Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu 1 5 10 15

Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
20 25 30

| Glu | Glu | Ala 35 | Tyr | Ile | Pro | Lys | Glu 40 | Gln | Lys | Tyr | Ser | Phe 45 | Leu | Gln | Asn |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Pro | Gln 50 | Thr | Ser | Leu | Çyş | Phe 55 | Ser | Glu | Ser | Ile | Pro 60 | Thr | Pro | Ser | Asn |
| Arg 65 | Glu | Glu | Thr | Gln | Gln 70. | Lys | Ser | Asn | Leu | Glu 75 | Leu | Leu | Arg | Ile | Ser 80 |
| Leu | Leu | Leu | Ile | Gln 85 | Ser | Trp | Leu | Glu | Pro 90 | Val | Gln | Phe | Leu | Arg 95 | Ser |
| Val | Pḥe | Ala | Asn 100 | Ser | Leu | Val | Tyr | Gly 105 | Ala | Ser | Asp | Ser | Asn 110 | Val | Tyr |
| Asp | Leu | Leu 115 | ГÀг | Asp | Leu | Glu | Glu 120 | Gly | Ile | Gln | Thr | Leu 125 | Met | Gly | Arg |
| Leu | Glu 130 | Asp | Gly | Ser | Pro | Arg 135 | Thr | Gly | Gln | Ile | Phe 140 | ГÀЗ | Gln | Thr | Tyr |
| Ser 145 | Lys | Phe | Asp. | Thr | Asn 150 | Ser | His | Asn | Asp | Asp 155 | Ala | Leu | Leu | Lys | Asn 160 |
| Tyr | Gly | Leu | Leu | Туг 165 | Cys | Phe | Arg | Lys | Asp 170 | | Asp | Lys | Val | Glu 175 | Thr |
| Phe | Leu | Arg | Ile 180 | Val | Gln | Cys | Arg | Ser 185 | Val | · Glu | Gly | Ser | Cys 190 | | Phe |
| Phe | Pro | Thr 195 | Ile | Pro | Leu | Ser | Arg 200 | Leu | Phe | Asp | Asn | Ala 205 | Met | Leu | Arg |
| Ala | His 210 | Arg | Leu | His | Gln | Leu 215 | Ala | Phe | Asp | Thr | Tyr 220 | Gln | Glu | Phe | Glu |
| Glu 225 | Ala | Tyr | Ile | Pro | Lys 230 | Glu | Gln | Lys | Tyr | Ser 235 | | Leu | Gln | Asn | Pro 240 |
| Gln | Thr | Ser | Leu | Cys 245 | Phe | Ser | Glu | Ser | Ile 250 | Pro | Thr | Pro | Ser | Asn 255 | Arg |
| Glu | Glu | Thr | Gln 260 | | Lys | Ser | Asn | Leu 265 | | Leu | Leu | Arg | 11e 270 | Ser | Leu |
| Leu : | Leu | 11e 275 | | Ser | Trp | Leu | Glu 280 | | Val | Gln | Phe | Leu 285 | | Ser | Val |
| Phe | Ala 290 | | Ser | Leu | Val | Tyr 295 | | Ala | Ser | Asp | Ser 300 | | Val | Tyr | `Asp |
| Leu 305 | | Lys | Asp | Leu | Glu 310 | | Gly | Ile | Gln | Thr 315 | | . Met | Gly | Arg | Leu 320 |
| Glu | Asp | Gly | Ser | Pro 325 | | Thr | Gly | Gln | Ile 330 | | Lys | Gln | Thr | Tyr 335 | Ser |

- Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr 340 345 350
- Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe 355 360 365
- Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Phe 370 375 380
- Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala 385 390 395 400
- His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu 405 410 415
- Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln
 420 425 430
- Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu
 435
 440
 445
- Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu Leu 450 455 460
- Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe 465 470 475 480
- Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu 485 490 495
- Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu Glu 500 505 510
- Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys 515 520 525
- Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly 530 540
- Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu 545 550 555 560
- Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe 565 570
- <210> 33
- <211> 55
- <212> DNA
- <213> Artificial
- <220>
- <223> synthetic sequence
- <400> 33
- taccatatga catgatcatg tggcttcggt ttcccaacca ttcccttatc cagge

| <210 <211 <212 <213 | > ! > I | 34 591 ONA Artii | ficia | ıl | | | | | | | | | | | | |
|------------------------------|------------|---------------------------|-----------|----------|-------|-----------|-----------|-----------|---------------|------|-----------|-----------|-------|-----------|--------|--|
| <220 <223 | | syntl | netic | : seq | luenc | :e | | | | | | | | | | |
| <400 cata | | 34 cat q | gatca | ıtgtg | ıg ct | tegg | jtttc | cca | acca | ittc | cctt | atco | ag g | gcttt | ttgac | |
| aacg | cta | tgc 1 | tccgc | :gccc | a to | gtct | gcac | caç | ctgg | jcct | ttga | caco | ta (| ccagg | agttt | |
| gaag | aag | cct a | atato | ccaa | ia gg | jaaca | igaag | tat | tcat | tcc | tgca | gaac | cc (| ccaga | cctcc | |
| ctct | gtt | tct (| cagag | tcta | it to | cgac | caccc | tcc | aaca | rggg | agga | aaca | ica a | acaga | aatcc | |
| aacc | tag | agc 1 | tgctc | ecgea | ıt ct | ccct | gctg | cto | atco | agt | cgtg | ggctg | gga g | gadag | ıtgcag | |
| ttcc | tca | gga g | gtgto | ttcg | jc ca | acag | gcctg | gto | jtac <u>c</u> | gcg | ccto | etgac | ag | caacg | ıtctat | |
| gaco | tcc | taa a | aggac | ctag | ja gg | jaagg | gcatc | caa | acgo | tga | tggg | ggagg | jct (| ggaag | jatggc | |
| agec | CCC | gga (| ctggg | gcaga | t ct | tcaa | agcag | aco | taca | igca | agtt | cgad | cac. | aaact | cacac | |
| aacg | atg | acg (| cacta | ctca | a ga | acta | cggg | cto | jctct | act | gctt | cago | gaa (| ggaca | tggac | |
| aagg | jtcg. | aga (| catto | ctgo | g ca | tcgt | gcag | tgo | cgct | ctg | tgga | aggga | atc. | С | | |
| <210 <211 <212 <213 | .> !> | 35 192 PRT Arti: | ficia | ıl | | | | | | | . , | | | | | |
| <220 <223 | | synt | hetio | c sec | ueno | e | | | , | | | | | | | |
| <400 | | 35 | | | - | | | | | | | | | | | |
| Ser 1 | Cys | Gly | Phe | Gly 5 | Phe | Pro | Thr | Ile | Pro 10 | Leu | Ser | Arg | Leu | Phe 15 | Asp | |
| Asn | Ala | Met | Leu 20 | Arg | Ala | His | Arg | Leu 25 | His | Gln | Leu | Ala | Phe | Asp | Thr | |
| Tyr | Gln | Glu 35 | Phe | Glų | Glu | Ala | Tyr 40 | Ile | Pro | Lys | Glu | Gln 45 | Lys | Tyr | Ser | |
| Phe | Leu 50 | Gln | Asn | Pro | Gln | Thr 55 | Ser | Leu | Cys | Phe | Ser 60 | Glu | Ser | Ile | Pro | |

Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu

Leu Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln

95

Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp 100 105 Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr 120 Leu Met Gly Arq Leu Glu Asp Gly Ser Pro Arq Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala 150 155 Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly 185 <210> 36 <211> 1158 DNA · <212> <213> Artificial <220> <223> synthetic sequence <220> <221> misc_feature <222> (577)..(1152) sequence is repeated N-1 times, where N is a positive whole numbe <223> <400> 36 tgatcatgtg gcttcggttt cccaaccatt cccttatcca ggctttttga caacgctatg cteegegeee ategtetgea eeagetggee tttgacaeet accaggagtt tgaagaagee 120 tatateceaa aggaacagaa gtatteatte etgeagaace eccagacete cetetgttte 180 tcagagtcta ttccgacacc ctccaacagg gaggaaacac aacagaaatc caacctagag 240 etgeteegea tetecetget geteateeag tegtggetgg ageeegtgea gtteeteagg 300 agtgtcttcg ccaacagcct ggtgtacggc gcctctgaca gcaacgtcta tgacctccta 360 420 aaggacctag aggaaggcat ccaaacgctg atggggaggc tggaagatgg cagccccgg actgggcaga tetteaagea gaeetacage aagttegaea caaacteaca caaegatgae 480 gcactactca agaactacgg gctgctctac tgcttcagga aggacatgga caaggtcgag 540 acattectge geategtgea gtgeegetet gtggagggat eatgtggett eggttteeea 600

accattecet tatedagget ttttgacaac getatgetee gegeecateg tetgeaccag

660

| | ctggcctttg | acacctacca | ggagtttgaa | gaagcctata | tcccaaagga | açagaagtat | 720 |
|---|-----------------|------------|------------|------------|------------|------------|--------|
| | tcattcctgc | agaaccccca | gacctccctc | tgtttctcag | agtctattcc | gacaccctcc | 780 |
| | aacagggagg | aaacacaaca | gaaatccaac | ctagagctgc | tccgcatctc | cctgctgctc | 840 |
| | atccagtcgt | ggctggagcc | cgtgcagttc | ctcaggagtg | tcttcgccaa | cagcctggtg | 900 |
| | tacggcgcct | ctgacagcaa | cgtctatgac | ctcctaaagg | acctagagga | aggcatccaa | 960 |
| | acgctgatgg | ggaggctgga | agatggcagc | ccccggactg | ggcagatctt | caagcagacc | 1020 |
| | tacagcaagt : | tcgacacaaa | ctcacacaac | gatgacgcac | tactcaagaa | ctacgggctg | 1080 |
| - | ctctactgct | tcaggaagga | catggacaag | gtcgagacat | tcctgcgcat | cgtgcagtgc | 1140 |
| | cgctctgtgg | agggatcc | | | - | | , 1158 |

<211> 384

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> MISC_FEATURE

<222> (192)..(383)

<223> sequence is repeated N-1 times, where N is a positive whole numbe

<400> 37

Ser Cys Gly Phe Gly Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr 20 25 30

Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser 35 40 45

Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro 50 55 60

Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu
65 70 75 80

Leu Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln
85 90 95

Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp 100 105 110 Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr
115 120 125

Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe

Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala 145 150 155 160

Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp 165 170 175

Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly
180 185 190

Ser Cys Gly Phe Gly Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp 195 200 205

Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr 210 215 220

Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser 225 230 235 240

Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro 245 250 255

Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu 260 265 270

Leu Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln 275 280 285

Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp 290 295 300

Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr 305 310 315 320

Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe 325 330 335

Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala 340 345 350

Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp 355 360 365

Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly 370 375 380

<210> 38

<211> 1743

<212> DNA

<213> Artificial

<220>
<223> synthetic sequence
<220>
<221> misc_feature
<222> (1141)..(1716)
<223> sequence is repeated N-1 times, where N is a positive whole numbe

<400> 38 60 catatgttcc caaccattcc cttatccagg ctttttgaca acgetatgct ccgcgcccat cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag 120 gaacagaagt atteatteet geagaaceee cagaceteee tetgtttete agagtetatt 180 ecgacaccet ccaacaggga ggaaacacaa cagaaateca acetagaget getecgeate 240 tecetgetge teatecagte gtggetggag eeegtgeagt teeteaggag tgtettegee 300 aacagootgg tgtacggogo ototgacago aacgtotatg acotootaaa ggacotagag . 360 420 gaaggcatcc aaacgctgat ggggaggctg gaagatggca gccccggac tgggcagatc ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 480 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 540 ategtgeagt geegetetgt ggagggatea tgtggetteg gttteecaac cattecetta 600 tecaggettt ttgacaaege tatgeteege geecategte tgeaceaget ggeetttgae 660 acctaccagg agtttgaaga agcctatatc ccaaaggaac agaagtattc attcctgcag 720 aacccccaga cctccctctg tttctcagag tctattccga caccctccaa cagggaggaa 780 840 acacaacaga aatccaacct agaqctgctc cgcatctccc tgctgctcat ccagtcgtgg ctggagcccg tgcagttcct caggagtgtc ttcgccaaca gcctggtgta cggcgcctct 900 gacagcaacg tetatgacet cetaaaggae etagaggaag geatecaaae getgatgggg 960 aggetggaag atggeageee ceggaetggg cagatettea ageagaeeta cageaagtte 1020 gacacaaact cacacaacga tgacgcacta ctcaagaact acgggctgct ctactgcttc 1080 aggaaggaca tggacaaggt cgagacattc ctgcgcatcg tgcagtgccg ctctgtggag 1140 1200 ggatcatgtg gcttcggttt cccaaccatt cccttatcca ggctttttga caacgctatg 1260 ctccgcgccc atcgtctgca ccagctggcc tttgacacct accaggagtt tgaagaagcc tatatcccaa aggaacagaa gtattcattc ctgcagaacc cccagacctc cctctgtttc 1320 1380 tcagagteta tteegacaee etecaacagg gaggaaacae aacagaaate caacetagag

```
etgeteegea tetecetget geteateeag tegtggetgg agecegtgea gtteeteagg
agtgtcttcg ccaacagcct ggtgtacggc gcctctgaca gcaacqtcta tgacctccta .
                                                                    1500
aaggacctag aggaaggcat ccaaacgctg atggggaggc tggaagatgg cagccccgg
                                                                    1560
actgggcaga tcttcaagca gacctacagc aagttcgaca caaactcaca caacgatgac
                                                                    1620
gcactactca agaactacgg gctgctctac tgcttcagga aggacatgga caaggtcgag
                                                                    1680
acattectge geategtgea gtgeegetet gtggagggat eatgtggett etagtaggte
                                                                    1740
gac
                                                                    1743
<210>
       39
<211>
       576
<212>
       PRT
<213>
       Artificial
<220>
       synthetic sequence
<223>
<220>
      MISC FEATURE
<221>
<222>
       (380)..(571)
<223>
       sequence is repeated N-1 times, where N is a positive whole numbe
<220>
<221>
      mat peptide
<222>
       (1)..()
<400>
Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu
Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn
Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn
Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser
Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser
Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
            100
                                                    110
```

| Asp | Leu | Leu 115 | Lys | Asp | Leu | Glu | Glu 120 | Gly | Ile | Gln | Thr | Leu 125 | Met | Gly | Arg |
|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Leu | Glu 130 | Asp | Gly | Ser | Pro | Arg 135 | Thr | Gly | Gln | Ilė | Phe 140 | Lys | Gln | Thr | Tyr |
| Ser 145 | Lys | Phe | Asp | Thr | Asn 150 | Ser | His | Asn | Asp | Asp 155 | Ala | Leu | Leu | Lys | Asn 160 |
| Tyr | Gly | Leu | Leu | Tyr 165 | Cys | Phe | Arg | Lys | Asp 170 | Met | Asp | Lys | Val | Glu 175 | Thr |
| Phe | Leu | Arg | Ile 180 | Vạl | Gln | Cys | Arg | Ser 185 | Val | Glu | Gly | Ser | Cys 190 | Gly | Phe |
| Gly | Phe | Pro 195 | Thr | Ile | Pro | Leu | Ser 200 | Arg | Leu | Phe | Asp | Asn 205 | Ala | Met | Leu |
| Arg | Ala 210 | His | Arg | Leu | His | Gln 215 | Leu | Ala | Phe | Asp | Thr 220 | Tyr | Gln | Glu | Phe |
| Glu 225 | Glu | Ala | Tyr | Ile | Pro 230 | Lys | Glu | Gln | Lys | Tyr 235 | Ser | Phe | Leu | Gln | Asn 240 |
| Pro | Gln | Thr | Ser | Leu. 245 | Cys | Phe | Ser | Glu | Ser 250 | Ile | Pro | Thr | Pro | Ser 255 | Asn |
| Arg | Glu | Glu | Thr 260 | Gln | Gln | Lys | Ser | Asn 265 | Leu | Glu | Leu | Leu | Arg 270 | Ile | Ser |
| Leu | Leu | Leu 275 | Ile | Gln | Ser | Trp | Leu 280 | Glu | Pro | Val | Ģln | Phe 285 | Leu | Arg | Ser |
| Val | Phe 290 | Ala | Asn | Ser | Leu | Val 295 | Tyr | Gly | Ala | Ser | Asp 300 | | Asn | Val | Tyr |
| Asp 305 | Leu | Leu | Lys | Asp | Leu 310 | Glu | Glu | Gly | Ile | Gln 315 | | Leu | Met | Gly | Arg 320 |
| Leu | Glu | Asp | Gly · | Ser 325 | Pro | Arg | Thr | Gly | Gln 330 | Ile | Phe | Lys | Gln | Thr 335 | Tyr |
| Ser | Lys | Phe | Asp 340 | Thr | Asn | Ser | His | Asn 345 | Asp | Asp | Ala | Leu | Leu 350 | Lys | Asn |
| Tyr | Gly | Leu 355 | Leu | Tyr | Cys | Phe | Arg 360 | Lys | Asp | Met | Asp | Lys 365 | Val | Glu | Thr |
| Phe | Leu 370 | Arg | Ile | Val | Gln | Cys 375 | Arg | Ser | Val | Glu | Gly 380 | Ser | Сув | Gly | Phe |
| Gly 385 | Phe | Pro | Thr | Ile | Pro 390 | Leu | Ser | Arg | Leu | Phe 395 | Asp | Asn | Ala | Met | Leu 400 |
| Arg | Ala | His | Arg | Leu 405 | His | Gln | Leu | | Phe 410 | Asp | Thr | Tyr | Gln | Glu 415 | Phe |

| | | | | | | | | • | | | | | | | |
|---------------------------|-------------------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|-----|
| | | | | | | | | | | | | | • | | |
| Glu Gl | lu Ala | Tyr 420 | Ile | Pro | Lys | Gľu | Gln 425 | Lys | Tyr | Ser | Phe | Leu 430 | Gln | Asn | |
| Pro Gl | n Thr 435 | | Leu | Cys | Phe | Ser 440 | Glu | Ser | Ile | Pro | Thr 445 | Pro | Ser | Asn | |
| Arg Gl | | Thr | Gln | Gln | Lys 455 | Ser | Asn | Leu | Glu | Leu 460 | Leu | Arg | Ile | Ser | |
| Leu Le 465 | eu Leu | Ile | Gln | Ser 470 | Trp | Leu | Glu | Pro | Val 475 | | Phẹ | Leu | Arg | Ser 480 | |
| Val Ph | ne Ala | Asn | Ser 485 | Leu | Val | Tyr | Gly | Ala 490 | Ser | Asp | Ser | Asņ | Val 495 | Tyr | |
| Asp Le | eu Leu | Lys 500 | Asp | Leu | Glu | Glu | Gly. 505 | Ile | Gln | Thr | Leu | Met 510 | Gly | Arg | |
| Leu Gl | lu Asp 515 | | Ser | Pro | Arg | Thr 520 | Gly | Gln | Ile | Phe | Lys 525 | Gln | Thr | Tyr | |
| Ser Ly 53 | s Phe | Asp | Thr | Asn | Ser 535 | His | Asn | Asp | Asp | Ala 540 | Leu | Leu | Lys | Asn | • |
| Tyr Gl | ly Leu | Leu | Tyr | Cys 550 | Phe | Arg | Lys | Asp | Met 555 | Asp | Lys | Val | Glu | Thr 560 | |
| Phe Le | eu Arg | Ile | Val 565 | Gln | Cys | Arg | Ser | Val 570 | Glu | Gly | Ser | Cys | Gly 575 | Phe | |
| <210><211><211><212><213> | 40 39 DNA Arti | ficia | al | | | | · | | | • | | | | : | |
| <220> <223> | synt | heti | c se | quen | ce | | - | | | | | | | | |
| <400> | 40 | | | | | | | | | | | • | | | |
| cgcgga | atcct | catg | agaa | ge ca | acago | ctgc | c cto | ccaca | aga | · | | | | | 39 |
| <210><211><211> | 41 591 DNA | | | | | ; · | | | | | | | | | |
| <213> | | fici | al | | | | | | | | | | | | |
| <220> <223> | synt | heti | c șe | quen | ce | | | | | | | | | | |
| <400> | 41 gttcc | caac | catto | cc c1 | ttato | ccag | g çti | cttt | gaca | acg | ctate | gct (| ccgc | gcccat | 60 |
| cgtctg | gçacc | agct | ggcc | tt t | gacad | ccta | c cag | ggagt | tttg | aag | aagc | cta 1 | tatc | ccaaag | 120 |
| gaacag | gaagt | attc | attc | ct go | : cagaa | accc | c cac | gacci | taca | tct | gttt | ctc a | agag | tctatt | 180 |

HC (F

| ccgacaccct | ccaacaggga | ggaaacacaa | cagaaatcca | acctagagct | gctccgcatc | 240 |
|------------|------------|------------|------------|------------|------------|-----|
| tccctgctgc | tcatccagtc | gtggctggag | cccgtgcagt | tcctcaggag | tgtcttcgcc | 300 |
| aacagcctgg | tgtacggcgc | ctctgacagc | aacgtctatg | acctcctaaa | ggacctagag | 360 |
| gaaggcatcc | aaacgctgat | ggggaggctg | gaagatggca | gcccccggac | tgggcagatc | 420 |
| ttcaagcaga | cctacagcaa | gttcgacaca | aactcacaca | acgatgacgc | actactcaag | 480 |
| aactacgggc | tgctctactg | cttcaggaag | gacatggaca | aggtcgagac | attcctgcgc | 540 |
| atcgtgcagt | gccgctctgt | ggagggcagc | tgtggcttct | catgaggatc | c . | 591 |
| | | | | | | |

<211> 193

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<400> 42

Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu 1 5 10 15

Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
20 25 30

Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn 35 40 45

Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn 50 55 60

Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser 65 70 75 80

Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser 85 90 95

Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
100 105 110

Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg
115 120 125

Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr
130 140

Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr 165 170 175

185 180 Ser 43 <210> <211> 50 <212> DNA Artificial <213> <220> <223> synthetic sequence <400> catgccatgg ggtggtggag gaagtttccc aaccattccc ttatccaggc 50 <210> 44 <211> 606 <212> DNA Artificial <213> <220> <223> synthetic sequence <400> 44 ccatggggtg gtggaggaag tttcccaacc attcccttat ccaggctttt tgacaacgct atgeteegeg eccategiet geaccagetg geetttgaca ectaccagga gtttgaagaa 120 gectatatee caaaggaaca gaagtattea tteetgeaga acceeeagae etecetetgt 180 ttctcagagt ctattccgac accctccaac agggaggaaa cacaacagaa atccaaccta 240 gagetgetee geateteest getgeteate cagtegtgge tggageeegt geagtteete 300 aggagtgtet tegecaacag cetggtgtae ggegeetetg acageaacgt etatgacete .360 ctaaaggacc tagaggaagg catccaaacg ctgatgggga ggctggaaga tggcagcccc 420 cggactgggc agatetteaa geagacetae ageaagtteg acacaaacte acacaacgat 480 gacgcactac tcaagaacta cgggctgctc tactgcttca ggaaggacat ggacaaggtc 540 gagacattcc tgcgcatcgt gcagtgccgc tctgtggagg gcagctgtgg cttctcatga 600 606 ggatcc <210> 45 198 <211> <212> PRT

Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe

<220>

<213>

Artificial

| <400 |)> 4 | 45 | | | | | | | | | | | | | | | |
|--------------------------|------------|--------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---|-----|
| Trp 1 | Gly | Gly | Gly | Gly 5 | Ser | Phe | Pro | Thr | Ile 10 | Pro | Leu | Ser | Arg | Leu 15 | Phe | | |
| Asp | Asn | Ala | Met 20 | Leu | Arg | Ala | His | Arg 25 | Leu | His | Gln | Leu | Ala 30 | Phe | Asp | | |
| Thr | Tyr | Gln 35 | Glu | Phe | Glu | Glu | Ala 40 | Tyr | Ile | Pro | Lys | Glu 45 | Gln | Lys | Tyr | | |
| Ser | Phe 50 | Leu | Gln | Asn | Pro | Gln 55 | Thr | Ser | Leu | Cys | Phe 60 | Ser | Glu | Ser | Ile | | |
| Pro 65 | Thr | Pro | Ser | Asn | Arg 70 | Glu | Glu | Thr | Gln | Gln 75 | Lys | Ser | Asn | Leu | Glu 80 | , | |
| Leu | Leu | Arg | Ile | Ser 85 | Leu | Leu | Leu | Ile | Gln 90 | Ser | Trp | Leu | Glu | Pro 95 | Val | | |
| Gln | Phe | Leu | Arg 100 | Ser | Val | Phe | Ala | Asn 105 | Ser | Leu | Val | Tyr | Gly 110 | Ala | Ser | | |
| Asp | Ser | Asn 115 | Val | Tyr | Asp | Leu | Lеџ 120 | Lys | Asp | Leu | Glu | Glu 125 | Gly | Ile | Gln | | |
| Thr | Leu 130 | Met | Gly | Arg | Leu | Gļu 135 | Asp | Gly | Ser | Pro | Arg 140 | Thr | Gly | Gln | Ile | | |
| Phe 145 | Lys | Gln | Thr | Tyr | Ser 150 | Lys | Phe | Asp | Thr | Asn 155 | Ser | His | Asn | Asp | Asp 160 | | |
| Ala | Leu | Leu | Lys | Asn 165 | Tyr | Gly | Leu | Ļeu | Tyr 170 | Cys | Phe | Arg | Lys | Asp 175 | Met | | |
| Asp | Lys | Val | Glu 180 | Thr | Phe | Leu | Arg | Ile 185 | Val | Gln | Cys | Arg | Ser 190 | Val | Glu | | |
| Gly | Ser | Cys 195 | _ | Phe | Ser | | | ÷ | | | | | | | | | |
| <21 <21 <21 <21 | 1,> 2 > | 46 603 DNA Arti | fici | al | | | | ; | | | | | | | | | |
| <22 <22 | | synt? | heti | c se | quen | ce | | | ٠ | | | | | | | | |
| <40 cca | | 46 gtg | gtgg | agga | ag t | ttcc | caac | c at | taca | ttat | cca | ggct | ttt | tgac | aacgct | : | 60 |
| atg | ctcc | gcg | ccca | tcgt | ct g | cacc | agct | g gc | cttt | gaca | cct | acca | gga (| gttt | gaagaa | ì | 120 |
| gcc. | tata | tcc | caaa | ggaa | ca g | aagt | attc | a tto | cċtg | caga | acc | ccca | gac | ctcc | ctctgt | ; | 180 |

<223> synthetic sequence

| ttctcagagt | ctattccgac | accctccaac | agggaggaaa | cacaacagaa | atccaaccta | 240 |
|------------|------------|------------|------------|------------|------------|-----|
| gagetgetee | gcatctccct | gctgctcatc | cagtcgtggc | tggagcccgt | gcagttcctc | 300 |
| aggagtgtct | tcgccaacag | cctggtgtac | ggcgcctctg | acagcaacgt | ctatgacctc | 360 |
| ctaaaggacc | tagaggaagg | catccaaacg | ctgatgggga | ggctggaaga | tggcagcccc | 420 |
| cggactgggc | agatcttcaa | gcagacctac | agcaagttcg | acacaaactc | acacaacgat | 480 |
| gacgcactac | tcaagaacta | cgggctgctc | tactgcttca | ggaaggacat | ggacaaggtc | 540 |
| gagacattcc | tgcgcatcgt | gcagtgccgc | tctgtggagg | gcagctgtgg | cttctaggga | 600 |
| tcc | | | | | | 603 |

<211> 197

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<400> 47

Trp Gly Gly Gly Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe
1 5 10 15

Asp Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp 20 25 30

Thr Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr 35 40 45

Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile 50 55 60

Pro Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu 65 70 75 80

Leu Leu Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val 85 90 95

Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser 100 105 110

Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln
115 120 125

Thr Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile 130 135 140

Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp 145 150 155 160

Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met
165 170 175

Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu 180 185 190

Gly Ser Cys Gly Phe 195

<210> 48

<211> 1200

<212> DNA

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> misc feature

<222> (595)..(1188)

<223> sequence is repeated N-1 times, where N is a positive whole numbe

<400> 48 60 ccatggggtg gtggaggaag tttcccaacc attcccttat ccaggetttt tgacaacget atgeteegeg cecategiet geaccagetg geettigaca eetaccagga gittigaagaa 120 gectatatec caaaggaaca gaagtattea tteetgeaga acceecagae etcectetgt 180 ttctcagagt ctattccgac accetecaae agggaggaaa cacaacagaa atccaaceta 240 gagetgetee geateteeet getgeteate cagtegtgge tggageeegt geagtteete 300 aggagtgtet tegecaacag cetggtgtac ggegeetetg acageaacgt etatgacete 360 ctaaaggacc tagaggaagg catccaaacg ctgatgggga ggctggaaga tggcagccc 420 480 eggactggge agatetteaa geagacetae ageaagtteg acacaaacte acacaacgat gaegeactae teaagaaeta egggetgete taetgettea ggaaggaeat ggaeaaggte 540 gagacattee tgegeategt geagtgeege tetgtggagg geagetgtgg etteteatgg 600 ggtggtggag gaagtttccc aaccattccc ttatccaggc tttttgacaa cgctatgctc 660 egegeecate gtetgeacca getggeettt gaeacetace aggagtttga agaageetat 720 780 atcccaaagg aacagaagta ttcattcctg cagaaccccc agacctccct ctgtttctca gagtetatte egacaceete caacagggag gaaacacaac agaaatecaa eetagagetg 840 etcegeatet ceetgetget catecagteg tggetggage cegtgeagtt cetcaggagt 900 gtottogoca acagootggt gtacggogoc totgacagoa acgtotatga cotoctaaag 960 gacctagagg aaggcatcca aacgctgatg gggaggctgg aagatggcag cccccggact gggcagatct tcaagcagac ctacagcaag ttcgacacaa actcacacaa cgatgacgca 1080 ctactcaaga actacgggct gctctactgc ttcaggaagg acatggacaa ggtcgagaca 1140 1200 ttcctgcgca tcgtgcagtg ccgctctgtg gagggcaget gtggcttctc atgaggatcc <210> 49 <211> 396 <212> PRT <213> Artificial <220> <223> synthetic sequence <220> <221> MISC FEATURE <222> (198)..(395) <223> sequence is repeated N-1 times, where N is a positive whole numbe <400> 49 Trp Gly Gly Gly Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile 55 Pro Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu 75 70 Leu Leu Arg Ile Ser Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser 105 Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln 115 Thr Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile 135 140 Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp

155

150

| Ala | Leu | Leu | Lys | Asn 165 | Tyr | Gly | Leu | Leu | Tyr 170 | Cys | Phe | Arg | Lys | Asp 175 | Met |
|------------------------------|--------------|----------------------------|------------|------------|------------|------------|------------|------------|------------|--------------------|------------|------------|------------|------------|------------|
| Asp | Lys | Val | Glu 180 | Thr | Phe | Leu | Arg | Ile 185 | Val | Gln | Cys | Arg | Ser 190 | Val | Glu |
| Gly | Ser | Cys 195 | Gly | Phe | Ser | Trp | Gly 200 | Gly | Gly | Gly | Ser | Phe 205 | Pro | Thr | Ile |
| Pro | Leu 210 | Ser | Arg | Leu | | Asp 215 | Asn | Ala | Met | Leu | Arg 220 | Ala | His | Arg | Leu |
| His 225 | Gln | Leu | Ala | Phe | Asp 230 | Thr | Tyr | Gln | Glu | Phe 235 | Glu | Glu | Ala | Tyr | Ile 240 |
| Pro | ГÀЗ | Glu | Gln | Lys 245 | Tyr | Ser | Phe | Leu | Gln 250 | Asn | Pro | Gln | Thr | Ser 255 | Leu |
| Cys | Phe | Ser | Glu 260 | Ser | Ile | Pro | Thr | Pro 265 | Ser | Asn | Arg | Glu | Glu 270 | Thr | Gln |
| Gln | Lys | Ser 275 | Asn | Leu | Glu | Leu | Leu 280 | Arg | Ile , | Ser | Leu | Leu 285 | Leu | Ile | Gln |
| Ser | Trp 290 | Leu | Glu | Pro | Val | Gln 295 | Phe | Leu | Arg | Ser | Val 300 | | Ala | Asn | Ser |
| Leu 305 | Val | Tyr | Gly | Ala | Ser 310 | Asp | Ser | Asn | Val | Ту <u>г</u> 315 | Asp | Leu | Leu | Lys | Asp 320 |
| Leu | Glu | Glu | Gly | Ile 325 | Gln | Thr | Leu | Met | Gly 330 | Arg | Leu | Glu | Asp | Gly 335 | Ser |
| Pro | Arg | Thr | Gly 340 | Gln | Ile | Phe | Lys | Gln 345 | Thr | Tyr | Ser | Lys | Phe 350 | Asp | Thr |
| Asn | Ser | His 355 | Asn | Asp | Asp | Ala | Leu 360 | Leu | Lys | Asn | Tyr | Gly 365 | Leu | Leu | Tyr |
| Cys | Phe 370 | | Lys | Asp | | Asp 375 | _ | Val | Glu | Thr | Phe 380 | | Arg | Ile | Val |
| Gln 385 | Cys | Arg | Ser | Val | Glu 390 | Gly | Ser | Cys | Gly | Phe 395 | Ser | | | | |
| <210 <210 <210 <210 | 1> : 2> ! | 50 1185 DNA Arti: | ficia | al | | | | | | | | | | | |
| <22 <22 | | syntl | heti | c sed | quen | ce | | | | | | | | | |
| <40 | | 50 tcc (| caac | catt | cc c | ttaț | ccag | g ct | tttt | gaca | acg | ctat | gct (| ccgc | gcccat |
| çgt | ctgc | acc a | agct | ggcc | tt t | gaca | ccta | c ca | ggag | tttg | aag | aagc | cta | tato | ccaaag |

```
gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt
                                                                      180
ccgacaccct ccaacaggga ggaaacacaa cagaaatcca acctagagct gctccgcatc
                                                                     240
tecetgétge teatecagte qtqqctqqaq ceegtgeagt teeteaggag tgtettegee
                                                                      300
aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag
                                                                      360
gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc
                                                                    420
ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag
                                                                      480
aactacqqqc tqctctactq cttcaqqaaq qacatggaca aggtcgagac attcctgcgc
                                                                    540
                                                                      600
atogtgcagt googototgt ggagggcago tgtggcttot catggggtgg tggaggaagt
ttcccaacca ttcccttatc caqqcttttt qacaacqcta tgctccgcgc ccatcqtctg
                                                                      660
caccagetgg cetttgacac etaccaggag tttgaagaag cetatatece aaaggaacag
                                                                      720
aagtattcat teetgeagaa eeeccagaee teeetetgtt teteagagte tatteegaca
                                                                      780
ccctccaaca gggaggaaac acaacagaaa tccaacctag agctgctccg catctccctg
                                                                      840
etgeteatee agtegtgget ggageegtg eagtteetea ggagtgtett egeeaacage
                                                                      900
ctggtgtacg gcgcctctga cagcaacgtc tatgacctcc taaaggacct agaggaaggc
                                                                      960
atccaaacgc tgatggggag gctggaagat ggcagcccc ggactgggca gatcttcaag
                                                                     1020
caqaectaca qeaaqtteqa cacaaactea cacaacgatg acgeactact caagaactac
                                                                     1080
gggctgctct actgcttcag gaaggacatg gacaaggtcg agacattcct gcgcatcgtg
                                                                     1140
                                                                     1185
cagtgccgct ctgtggaggg cagctgtggc ttctcatgag gatcc
```

```
<210> 51
```

<220>

<223> synthetic sequence

<220>

<221> mat_peptide

<222> (1)..()

<400> 51

Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu 1 5 10 15

Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe

<211> 391

<212> PRT

<213> Artificial

| Glu | Glu | Ala 35 | Tyr | Ile | Pŗo | Lys | Glu 40 | Gln | Lys | Tyr | Ser | Phe 45 | Leu | Gln | Asn |
|------------|------------|-------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Pro | Gln 50 | Thr | Ser | Leu | Cys | Phe 55 | Ser | Glu | Ser | Ile | Pro 60 | Thr | Pro | Ser | Asn |
| Arg 65 | Glu | Glu | Thr | Gln | Gln 70 | Lys | Ser | Asn | Leu | Glu 75 | Leu | Leu | Arg | Iļe | Ser 80 |
| Leu | Leu | Leu | Ile | Gln 85 | Ser | Trp | Leu | | Pro 90 | Val | Gln | Phe | Leu | Arg 95 | Ser |
| Val | Phe | Ala | Asn 100 | Ser | Leu | Val | Tyr | Gly 105 | Ala | Ser | Asp | Ser | Asn 110 | Val | Tyr |
| Asp | Leu | Leu 115 | Lys | Asp | Leu | Glu | Glu 120 | Gly | Ile | Gln | Thr | Leu 125 | Met | Gly | Arg |
| Leu | Glu 130 | Asp | Ġly | Ser | Pro | Arg 135 | Thr | Gly | Gln | ·Ile | Phe 140 | Lys | Gln | Thr | Tyr |
| Ser 145 | Lys | Phe | Asp | Thr | Asn 150 | Ser | His | Asn | Asp | Asp 155 | Ala | Leu | Leu | Lys | Asn 160 |
| Tyr | Gly | Leu | Leu | Tyr 165 | Cys | Phe | Arg | Lys | Asp 170 | Met | Asp | Lys | Val | Glu 175 | |
| Phe | Leu | Arg | Ile 180 | Val | ·Gln | Cys | Arg | Ser 185 | Val | Glu | Gly | Ser | Cys 190 | Gly | Phe |
| Ser | Trp | .Gly 195 | Gly | Gly | Gly | Ser | Phe 200 | Pro | Thr | Ile | Pro | Leu 205 | Ser | Arg | Leu |
| Phe | Asp 210 | Asn | Ala | Met | Leu | Arg 215 | Ala | His | Arg | Leu | His 220 | Gln | Leu | Ala | Phe |
| Asp 225 | Thr | Tyr | Gln | Glu | Phe 230 | Glu | Glu | Ala | Tyr | Ile 235 | Pro | Lys | Glu | Gln | Lys 240 |
| Tyr | Ser | Phe | Leu | Gln 245 | Asn | Pro | Gln | Thr | Ser 250 | Leu | Cys | Phe | Ser | Glu 255 | Ser |
| Ile | Pro | Thr | Pro. 260 | Ser | Asn | Arg | Glu | Glu 265 | Thr | Gln | Gln | Г'ns | Ser 270 | Asn | Leu |
| Glu | Leu | Leu 275 | Arg | Ile | Ser | Leu | Leu 280 | Leu | Ile | Gln | Ser | Trp 285 | Leu | Glu | Pro |
| Val | Gln 290 | Phe | Leu | Arg | Ser | Val 295 | Phe | Ala | Asn | Ser | Leu 300 | Val | Tyr | Gly | Ala, |
| Ser 305 | Asp | Ser | Asn | Val | Tyr 310 | Asp | Leu | Leu | Lys | Asp 315 | Leu | Glu | Glu | Gly | Ile 320 |
| Gln | Thr | Leu | Met | Gly 325 | Arg | Leu | Glu | Asp | Gly 330 | Seŗ | Pro | Arg | Thr | Gly 335 | Gln |

Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp 340 Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Ser 385 390 <210> 52 <211> 1779 <212> DNA Artificial <213> <220> <223> synthetic sequence <220> <221> misc feature (1174)..(1767) <222> <223> sequence is repeated N-1 times, where N is a positive whole numbe <4.00>

catatgttcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgcccat .60 cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag 120 gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt 180 ccgacaccct ccaacaggga ggaaacacaa cagaaatcca acctagagct gctccgcatc 240 tecetgetge teatecagte gtggetggag ceegtgeagt teeteaggag tgtettegee 300 aacageetgg tgtaeggege etetgaeage aacgtetatg aceteetaaa ggaeetagag 360 gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 420 ttcaagcaga cetacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 480 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 540 atcgtgcagt gccgctctgt ggagggcagc tgtggcttct catggggtgg tggaggaagt 600 ttcccaacca ttcccttatc caggettttt gacaacgeta tgetcegege ccategtetg 660 caccagetgg cetttgacac ctaccaggag tttgaagaag cetatateec aaaggaacag 720 aagtattoat tootgoagaa coccoagaco tooctotgtt totoagagto tattoogaca 780 ccctccaaca gggaggaaac acaacagaaa tccaacctag agctgctccg catctccctg 840

```
etgeteatee agtegtgget ggageeegtg eagtteetea ggagtgtett egeeaacage
                                                                     900
ctggtgtacg gcgcctctga cagcaacgtc tatgacctcc taaaggacct agaggaaggc
                                                                     960
atccaaacgc tgatggggag gctggaagat ggcagcccc ggactgggca gatcttcaag
                                                                    1020
cagacctaca gcaagttcga cacaaactca cacaacgatg acgcactact caagaactac
                                                                    1080
gggetgetet actgetteag gaaggacatg gacaaggteg agacatteet gegeategtg
                                                                    1140
cagtgccgct ctgtggaggg cagctgtggc ttctcatggg gtggtggagg aagtttccca
                                                                    1200
accettecet tatecagget ttttgacaac getatgetee gegeegateg tetgeaccag
                                                                    1260
ctggcctttg acacctacca ggagtttgaa gaagcctata tcccaaagga acagaagtat
                                                                    1320
tcattcctgc agaaccccca gacctccctc tgtttctcag agtctattcc gacaccctcc
                                                                    1380
aacagggagg aaacacaaca gaaatccaac ctagagctgc tccgcatctc cctgctgctc
                                                                    1440
atccagtcgt ggctggagcc cgtgcagttc ctcaggagtg tcttcgccaa cagcctggtg
                                                                    1500
tacqqcqcct ctgacagcaa cgtctatgac ctcctaaaqq acctaqagga aggcatccaa
                                                                    1560 -
acgctgatgg ggaggctgga agatggcagc ccccggactg ggcagatctt caagcagacc
                                                                    1620
tacagcaagt tcgacacaaa ctcacacaac gatgacgcac tactcaagaa ctacgggctg
                                                                    1680
ctctactgct tcaggaagga catggacaag gtcgagacat tcctgcgcat cgtgcagtgc
                                                                    1740
cgctctgtgg agggcagctg tggcttctca tgaggatcc
                                                                    1779
```

```
<210> 53
<211> 589
```

<220>

<223> synthetic sequence

<220>

<221> MISC_FEATURE

<222> (391)..(588)

<223> sequence is repeated N-1 times, where N is a positive whole numbe

<220>

<221> mat peptide

<222> (1)..()

<400> 53

Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu 1 5 10 15

<212> PRT <213> Artificial

- Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe 20 25 30
- Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn 35 40 45
- Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn 50 55 60
- Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser 65 70 75 80
- Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser 85 90 95
- Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
 100 105 110
- Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg 115 120 125
- Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr 130 135 140
- Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr 165 170 175
- Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe 180 185 190
- Ser Trp Gly Gly Gly Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu 195 200 205
- Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe 210 215 220
- Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys 225 230 235 240
- Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser 245 250 255
- Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu 260 265 270
- Glu Leu Leu Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro 275 280 285
- Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala 290 295 300
- Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile 305 310 315 320

Gln Thr Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln 325 330 335

Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp 340 345 350

Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp 355 360 365

Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val 370 375 380

Glu Gly Ser Cys Gly Phe Ser Trp Gly Gly Gly Gly Ser Phe Pro Thr 385 390 395 400

Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg 405 410 415

Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr 420 425 430

Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser 435 440 445

Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr 450 455 460

Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu Leu Leu Ile 465 470 475 480

Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn 485 490 495

Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys 500 505 510

Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu Glu Asp Gly 515 520 525

Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp 530 535 540

Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu 545 550 555 560

Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile 565 570 575

Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Ser 580 585

<210> 54

<211> 2370

<212> DNA

<213> Artificial

<220>

<221> misc_feature

<222> (1174)..(1767)

<223> sequence is repeated N-1 times, where N is a positive whole numbe

<400> 54 catatgttcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgcccat cgtctgcacc agetggcctt tgacacctac caggagtttg aagaagccta tateccaaag 120 180 gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt ccgacaccct ccaacaggga ggaaacacaa cagaaatcca acctagagct gctccgcatc 240 tecetgetge teatecagte gtggetggag ceegtgeagt teeteaggag tgtettegee 300 aacageetgg tgtaeggege etetgaeage aacgtetatg accteetaaa ggaeetagag 360 gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 420 480 ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 540 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 600 ategtgcagt geogetetgt ggagggcage tgtggcttet catggggtgg tggaggaagt ttcccaacca ttcccttate caggettttt gacaacgeta tgctccgcgc ccatcgtctg 660 720 caccagetgg cetttgacac etaccaggag tttgaagaag cetatatece aaaggaacag aagtattcat teetgeagaa eeeccagaee teeetetgtt teteagagte tatteegaca 780 840 ccctccaaca gggaggaaac acaacagaaa tccaacctag agctgctccg catctccctg 900 etgeteatee agtegtgget ggageeegtg cagtteetea ggagtgtett egeeaacage ctggtgtacg gcgcctctga cagcaacgtc tatgacctcc taaaggacct agaggaaggc 960 atccaaacgc tgatggggag gctggaagat ggcagcccc ggactgggca gatcttcaag 1020 cagacctaca gcaagttcga cacaaactca cacaacgatg acgcactact caagaactac 1080 gggctgctct actgcttcag gaaggacatg gacaaggtcg agacattcct gcgcatcgtg 1140 cagtgccgct ctgtggaggg cagctgtggc ttctcatggg gtggtggagg aagtttccca 1200 accattccct tatccaggct ttttgacaac gctatgctcc gcgcccatcg tctgcaccag 1260 1320 ctggcctttg acacctacca ggagtttgaa gaagcctata tcccaaagga acagaagtat teatteetge agaaccecca gaccteette tgttteteag agtetattee gacaccetee 1380 aacagggagg aaacacaaca gaaatccaac ctagagctgc tccgcatctc cctgctgctc 1440

```
atccagtcgt ggctggagcc cgtgcagttc ctcaggagtg tcttcgccaa caqcctqqtq
tacggcgcct ctgacagcaa cgtctatgac ctcctaaagg acctagagga aggcatccaa
                                                                    1560
acgctgatgg ggaggctgga agatggcagc ccccggactg ggcagatctt caagcagacc
                                                                    1620
tacagcaagt tcgacacaaa ctcacacaac gatgacgcac tactcaagaa ctacgggctg
                                                                    1680
ctctactgct tcaggaagga catggacaag gtcgagacat tcctgcgcat cgtgcagtgc
                                                                    1740
cgctctgtgg agggcagctg tggcttctca tggggtggtg gaggaagttt cccaaccatt
                                                                    1800
ccettateca ggetttttga caacgetatg etcegegeec ategtetgea ccagetggee
                                                                    1860
tttgacacct accaggagtt tgaagaagcc tatatcccaa aggaacagaa gtattcattc
                                                                    1920
ctgcagaacc cccagacctc cctctgtttc tcagagtcta ttccgacacc ctccaacagg
                                                                    1980
gaggaaacac aacagaaatc caacctagag ctgctccgca tctccctgct gctcatccag
                                                                    2040
tegtggetgg agecegtgea gtteeteagg agtqtetteg ceaacageet ggtgtaegge
                                                                    2100
gcctctgaca gcaacgtcta tgacctccta aaggacctag aggaaggcat ccaaacgctg
                                                                    2160
atggggaggc tggaagatgg cagccccgg actgggcaga tettcaagca gacctacagc
                                                                    2220
aagttegaca caaactcaca caacgatgac gcactactca agaactacgg gctgctctac
                                                                    2280
tgcttcagga aggacatgga caaggtcgag acattcctgc gcatcgtgca gtgccgctct
                                                                    2340
gtggagggca gctgtggctt ctagggatcc
                                                                    2370
```

```
<210> 55
<211> 786
<212> PRT
<213> Artificial
```

<220>

<223> synthetic sequence

<220>

<221> MISC_FEATURE <222> (391)..(588)

<220>

<221> mat_peptide

<222> (1)..()

<400> 55

Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu 1 5 10 15

Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
20 25 30

Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser 70 75 Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg .115 120 Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn 150 155 Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe 185 Ser Trp Gly Gly Gly Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu 195 200 Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe 215 Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys 230 Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu 260 265 Glu Leu Leu Arg Ile Ser Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala 295 Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile 305 310 Gln Thr Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln

330

335

- Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp 340 345 350
- Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp 355 360 365
- Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val 370 375 380
- Glu Gly Ser Cys Gly Phe Ser Trp Gly Gly Gly Gly Ser Phe Pro Thr 385 390 395 400
- Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg
 405 410 415
- Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr
 420 425 430
- Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser 435 440 445
- Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr 450 455 460
- Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu Leu Leu Ile 465 470 480
- Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn 485 490 495
- Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys 500 505 510
- Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu Glu Asp Gly 515 520 525
- Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp 530 535 540
- Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu 545 550 555 560
- Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile 565 570 575
- Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Ser Trp Gly Gly
 580 585 590
- Gly Gly Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala 595 600 605
- Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln 610 615 620
- Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu 625 630 635 640

| Gln Ası | ı Pro | Gln | Thr 645 | Ser | Leu | Cys | Phe | Ser 650 | Glu | Ser | Ile | Pro | Thr 655 | Pro | | | |
|-----------------|--------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---|----|----|
| Ser Ası | n Arg | Glu 660 | Glu | Ţhr | Gln | Gln | Lys 665 | Ser | Asn | Leu | Glu | Leu 670 | Leu | Arg | | | |
| Ile Sei | 675 | Leu | Leu | Ile | Gln | Ser 680 | Trp | Leu | Glu | Pro | Val 685 | Gln | Phe | Leu | | | |
| Arg Ser | | Phe | Ala | Asn | Ser 695 | Leu | Val | Tyr | Gly | Ala 700 | Ser | Asp | Ser | Asn | | | |
| Val Tyr | r .Asp | Leu | Leu | Lys 710 | Asp | Leu | Glu | Glu | Gly 715 | Ile | Gln | Thr | Ļeu | Met 720 | | e: | |
| Gly Arg | g Leu | Glu | Asp 725 | Gly | Ser | Pro | Arg | Thr 730 | Gly | Gln | Ile | Phe | Lys 735 | Gln | | | |
| Thr Ty | r Ser | Lys 740 | Phe | Asp | Thr | Asn | Ser 745 | His | Asn | Asp | Asp | Ala 750 | Leu | Leu | | | |
| Lys Ası | n Tyr 755 | Gly | Leu | Leu | Tyr | Cys 760 | Phe | Arg | Lys | Asp | Met 765 | | Lys | Val | | | |
| Glu Th | | Leu | Arg | Ile | Val 775 | Gln | Cys | Arg | Ser | Val 780 | Glu | Gly | Ser | Cys | | | |
| Gly Pho 785 | e | | | | | | | | | | | | | | | | |
| <210> <211> | 56 33 | | | | ٠. | | • | | | | | ÷ | | | | | |
| <212> <213> | DNA Arti | ficia | al | | | | | | | | | ř. | | | • | | - |
| <220> <223> | syntl | hetio | c se | quen | ce | | | | | | | | | | | | |
| <400> ttacca | 56 tgga | ttgc | cggc | gg c | ggçg | gatc | c aa | t | | • | | , | | | | | 33 |
| <210> | 5 <i>7</i> | | • | | | | | | | | | | | • | | | |
| <211> <212> | 36 DNA | | | | | | | | | | | | | | | | |
| <213> | Arti | ficia | al | | | | | | | | | | | | | - | |
| <220> <223> | synt | heți | c se | quen | ce | | | | | | | ٠ | | | | | |
| <400> ttacca | 57 tgga | tttg | atca | gg c | ggcg | gcgg | a tc | caat | | | | | | | | | 36 |
| <210> <211> | 58 36 | | | | | | | | | | | | | | | | |
| | 20 | | | | | | | | | | | | | | | | |

```
<212> · DNA
<213> Artificial
<220>
<223> synthetic sequence
<400> 58
tgatcaggcg gcggcggatc aggcggcggc ggatcc
                                                                   36
<210> 59
<211> 10
<212> PRT
<213> Artificial
<220>
<223> synthetic sequence
<400> 59
Ser Gly Gly Gly Ser Gly Gly Gly
               5
<210> 60
     48
<211>
<212> DNA
<213> Artificial
<220>
<223> synthetic sequence
<400> 60
gccggcggcg gcggatcagg cggcggcgga tcaggcggcg gcggatcc
                                                                    48
<210>
      61
<211>
      14
<212> PRT
<213> Artificial
<220>
<223> synthetic sequence
<400> 61
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
<210> 62
<211> 43
<212> DNA
<213> Artificial
<220>
<223> synthetic sequence
<400> 62
```

| ggaçata | atgc tgtgatc | att cccaaco | att ccctt | atcca ggc | | | 43 |
|---------------------------|--------------------------------|--------------|------------|-------------|-----------|------------|-------------|
| <210><211><211><212><213> | 63 41 DNA Artificial | | | | | | |
| <220> <223> | synthetic s | sequence | | | | | |
| <400> cgcgaat | 63 ttcg atccatg | ggaa gccacag | etg cccto | ccacag a | | | 41 |
| <210><211><211> | 64 36 DNA | | | | • | · | |
| <213> | Artificial | | | | | | |
| <220> <223> | synthetic s | sequence | . • | | | | |
| <400> cgcgtc | 64 gacc tagaago | ccac agetge | ctc cacaç | ga | | | 36 |
| <210><211><211><212><213> | 65 602 DNA Artificial | | · | | | | - - - |
| <220> <223> | synthetic s | sequence | | | | • | · . |
| <400> catatg | 65 ctgt gatcatt | ccc aaccatt | ccc ttate | ccaggc tttt | tgacaa d | cgctatgctc | . 60 |
| cgcgcc | catc gtctgca | acca gctggco | ettt gacad | cctacc agga | igtttga a | agaagcctat | 120 |
| atccca | aagg aacagaa | agța ttcatto | cctg cagaa | accccc agac | ctccct o | ctgtttctca | 180 |
| gagtct | attc cgacacc | cctc caacag | ggag gaaad | cacaac agaa | atccaa (| cctagagctg | 240 |
| ctccgc | atct ccctgct | gct catccag | gtcg tggct | tggagc ccgt | gcagtt o | cctcaggagt | 300 |
| gtcttc | gcca acagcct | ggt gtacggo | egcc tctg | acagca acgt | ctatga d | cctcctaaag | 360 |
| gaccta | gagg aaggcat | cca aacgct | gatg gggag | gg¢tgg aaga | itggcág (| ccccggact | 420 |
| gggcag | atct tcaagca | agac ctacago | caag ttcga | acacaa acto | acacaa (| cgatgacgca | 480 |
| ctactc | aaga actacgg | ggct gctctad | tgc ttcag | ggaagg acat | ggacaa g | ggtcgagaca | 540 |
| ttcctg | cgca tcgtgca | agtg ccgctct | gtg gaggg | gcagct gtgç | jcttcca t | tggatcgaat | 600 |
| tc | | | | | | | 602 |

```
<210>
       66
<211>
       192
<212>
      PRT
<213>
      Artificial
<220>
<223>
      synthetic sequence
<400> 66
Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu
Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
                               25
Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn
Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn
Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser
                70
                                       75
Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser
Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
                               105
Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg
Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr
Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn
                   150
145
                                       155
Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr
Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe
                                185
<210>
       67
<211>
       600
<212>
      DNA
      Artificial
<213>
```

<220>

<223> synthetic sequence

<400> 67

| catatgctgt | gatcattccc | aaccattccc | ttatccaggc | tttttgacaa | cgctatgctc | 60 |
|------------|------------|------------|------------|------------|------------|------|
| cgcgcccatc | gtctgcacca | gctggccttt | gacacctacc | aggagtttga | agaagcctat | 120 |
| atcccaaagg | aacagaagta | ttcattcctg | cagaaccccc | agacctccct | ctgtttctca | 180 |
| gagtctattc | cgacaccctc | caacagggag | gaaacacaac | agaaatccaa | cctagagctg | 240 |
| ctccgcatct | ccctgctgct | catccagtcg | tggctggagc | ccgtgcagtt | cctcaggagt | 300 |
| gtcttcgcca | acagcctggt | gtacggcgcc | tctgacagca | acgtctatga | cctcctaaag | 360 |
| gacctagagg | aaggcatcca | aacgctgatg | gggaggctgg | aagatggcag | ccccggact | 420 |
| gggcagatet | tcaagcagác | ctacagcaag | ttcgacacaa | actcacacaa | cgatgacgca | .480 |
| ctactcaaga | actacgggct | gctctactgc | ttcaggaagg | acatggacaa | ggtcgagaca | 540 |
| ttcctgcgca | tcgtgcagtg | ccgctctgtg | gagggcagct | gtggcttcta | ggtcgacgcg | 600 |

<210> 68

<211> 192

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<400> 68

Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu 1 5 10 15

Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe 20 25 30

Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn 35 40 45

Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn 50 55 60

Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser 65 70 75 80

Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser 85 90 95

Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
100 105 110

Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg 115 120 125

Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr 130 135 140 Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe <210> 69 <211> 639 <212> DNA <213> Artificial <220> <223> synthetic sequence <400> catatgetgt gateatteec aaccatteec ttatecagge tttttgacaa egetatgete 60 egegeecate gtetgeacca getggeettt gacacetace aggagtttga agaageetat 120 ateccaaagg aacagaagta tteatteetg cagaaceeec agaceteect etgtttetea 180 gagtetatte egacaceete caacagggag gaaacacaac agaaatecaa eetagagetg 240 ctccgcatct ccctgctgct catccagtcg tggctggagc ccgtgcagtt cctcaggagt 300 qtetteqeea acageetggt gtaeggegee tetgacagea acqtetatga cetectaaag 360 gacctagagg aaggcatcca aacgctgatg gggaggctgg aagatggcag cccccggact 420 gggcagatet teaageagae etacageaag ttegacacaa aeteacacaa egatgaegea 480 ctactcaaga actacgggct gctctactgc ttcaggaagg acatggacaa ggtcgagaca 540 600 tteetgegea tegtgeagtg eegetetgtg gagggeaget gtggettegg eggeggegga tcaggcggcg gcggatcagg cggcggcgga tccgaattc 639 <210> 70 <211> 206 <212> PRT <213> Artificial <220> <223> synthetic sequence <400> 70 Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe

| Glu | Glu | Ala 35 | Tyr | Iļe | Pro | Lys | Glu 40 | Gln | Lys | Tyr | Ser | Phe 45 | Leu | Gln | Asn | | |
|------------------------------|------------|--------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|-----|
| Pro | Gln 50 | Thr | Ser | Leu | Cys | Phe 55 | Ser | Glu | Ser | Ile | Pro 60 | Thr | Pro | Ser | Asn | | |
| Arg 65 | Glu | Glu | Thr | Gln | Gln 70 | Lys | Ser | Asn | Leu | Glu 75 | Leu | Leu | Arg | Ile | Ser 80 | | |
| Leu | Leu | Leu | Ile | Gln 85 | Ser | Trp | Leu | Glu | Pro 90 | Val | Gln | Phe | Leu | Arg 95 | Ser | | |
| Val | Phe | Ala | Asn 100 | Ser | Leu | Val | Tyr | Gly 105 | Ala | Ser | Asp | Ser | Asn 110 | Val | Туŗ | | |
| Asp | Leu | Leu 115 | Lys | Asp | Leu | Glu | Glu 120 | Gly | Ile | Gln | Thr | Leu 125 | Met | Gly | Arg | | |
| Leu | Glu 130 | Asp | Gly | Ser | Pro | Arg 135 | Thr | Gly | Gln | Ile | Phe 140 | Lys | Gln | Thr | Tyr | 121 | |
| Ser 145 | Lys | Phe | Asp | Thr | Asn 150 | Ser | His | Asn | Asp | Asp 155 | Ala | Leu′ | Leu | Lys | Asn 160 | | |
| Tyr | Gly | Leu | Leu | Tyr 165 | Cys | Phe | Arg | Lys | Asp 170 | Met | Asp | Lys | Val | Glu 175 | Thr | : | |
| Phe | Leu | Arg | Ile 180 | Val | Gln | Cys | Arg | Ser 185 | Val | Glu | Gly | Ser | Cys 190 | Gly | Phe | | |
| Gly | Gly | Gly 195 | _ | Ser | Gly. | Gly | Gly 200 | Gly | Ser | Gly | Gly | Gly 205 | Gly | | | ٠ | |
| <210 <211 <211 <211 | L> 2> 1 | 71 630 DNA Arti | fici | a.1 | | | | | | | | | | | | | |
| <21. | | ALCI | LICI | d1 | | | | | | | | • | | | | | |
| <22 | | syntl | hetio | c sed | queno | ce | | | | | | | | | | | |
| <4.00 cata | | 71 tcc (| caac | catto | cc ct | tato | ccag | g cti | ttt | gaca | acgo | ctate | gct (| ccgc | gcccat | | 60 |
| cgt | ctgc | acc a | agct | ggect | t to | gacad | ccta | ç caç | ggagt | ttg | aaga | agco | cta 1 | tatco | ccaaag | 1 | 20 |
| gaad | caga | agt a | attca | attc | ct go | cagaa | accc | c cag | gacci | ccc | tct | gttt | ctc a | agagi | tctatt | 1 | .80 |
| ccga | acac | çct (| ccaa | cagg | ga gg | gaaa | caca | a caç | gaaal | cca | acci | aga | gct g | gctc | cgcatc | 2 | 40 |
| tcc | etge | tgc 1 | tcat | ccag | c gt | ggct | zgga | g cc | cgtg | cagt | teet | cago | gag 1 | tgtc | ttcgcc | . 3 | 00 |
| aaca | agcc | tgg 1 | tgtad | cggc | gc ct | ctg | acago | c aad | cgtc | atg | acçi | ccta | aaa g | ggac | ctagag | 3 | 60 |
| gaag | ggca | tcc a | aaac | gctga | at go | ggga | ggct | g gaa | agato | ggca | gcc | cccg | gac 1 | tggg | cagatc | 4 | 20 |

| ttcaagcaga cct | acagcaa gttcg | acaca aaçt | cacaca a | acgatgacgc a | actactdaag |
|------------------------|-----------------------|-----------------|------------------|--------------------|----------------|
| aactacgggc tgc | tctactg cttca | ggaag gaca | atggaca a | aggtcgagac a | attcetgege |
| atcgtgcagt gcc | gctctgt ggagg | gcage tgtg | ggcttcg (| gcggcggcgg a | atcaggcggc |
| ggcggatcag gcg | gcggcgg atccg | aattc | • | | |
| <210> 72 | • . | | | | |
| <211> 206 <212> PRT | | *, | | . | |
| <213> Artific | cial | | | | |
| <220> <223> synthet | ic sequence | | | | |
| <400> 72 | · - | | | | |
| Met Phe Pro Th | | _ | | Asp Asn Ala | |
| 1 | 5 | | 10 | : ml m 01 | 15 |
| Arg Ala His An | 7 | Leu Ala 1 25 | Pne Asp | inr lyr Gin 30 | Glu Pne |
| Glu Glu Ala Ty 35 | r Ile Pro Lys | Glu Gln 1 40 | Lys Tyr | Ser Phe Leu 45 | Gln Asn |
| Pro Gln Thr Se | er Leu Cys Phe 55 | Ser Glu | • | Pro Thr Pro 60 | Ser Asn |
| Arg Glu Glu Th | nr Gln Gln Lys 70 | Ser Asn l | Leu Glu 1 75 | Leu Leu Arg | Ile Ser 80 |
| Leu Leu Leu I | le Gln Ser Trp 85 | | Pro Val (90 | Gln Phe Leu | Arg Ser 95 |
| Val Phe Ala As | sn Ser Leu Val 00 | Tyr Gly 7 | Ala Ser i | Asp Ser Asn 110 | • |
| Asp Leu Leu Ly 115 | ys Asp Leu Glu | Glu Gly : | Ile Gln | Thr Leu Met 125 | Gly Arg |
| Leu Glu Asp G | ly Ser Pro Arg | | | Phe Lys Gln 140 | Thr Tyr |
| Ser Lys Phe As | sp Thr Asn Ser 150 | His Asn | Asp Asp . 155 | Ala Leu Leu | Lys Asn 160 |
| Tyr Gly Leu Lo | eu Tyr Cys Phe 165 | | Asp Met . 170 | Asp Lys Val | Glu Thr 175 |
| Phe Leu Arg I | le Val Gln Cys 30 | Arg Ser | Val Glu | Gly Ser Cys 190 | |
| Gly Gly Gly G 195 | ly Ser Gly Gly | Gly Gly 8 | Ser Gly | Gly Gly Gly 205 | |

```
<210>
       73
<211>
       1248
       DNA
<212>
       Artificial
<213>
<220>
<223>
       synthetic sequence
<220>
<221>
       misc feature
<222>
       (619) . . (1236)
```

<223>

sequence is repeated N-1 times, where N is a positive whole numbe

<400> 73 tgatcattcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgcccat 60 egtetgeace agetggeett tgacacetae caggagtttg aagaageeta tateecaaag .120 gaacagaagt attcattcct gcagaacccc cagacctccc tctgtttctc agagtctatt 180 ecgacaccet ccaacaggga ggaaacacaa cagaaateca acetagaget getecgeate 240 300 teeetgetge teatecagte gtggetggag ceegtgeagt teeteaggag tgtettegee aacagcctgg tgtacggcgc ctctgacagc aacgtctatg acctcctaaa ggacctagag 360 gaaggcatcc aaacgctgat ggggaggctg gaagatggca gcccccggac tgggcagatc 420 ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 480 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 540 ategtgeagt geegetetgt ggagggeage tgtggetteg geggeggegg ateaggegge 600 660 ggcggatcag gcggcggcgg atcattccca accattccct tatccaggct ttttgacaac getatgetee gegeeeateg tetgeaceag etggeetttg acacetacea ggagtttgaa 720 gaageetata teecaaagga acagaagtat teatteetge agaaceeeca gaeeteeete 780 tgtttctcag agtctattcc gacaccctcc aacagggagg aaacacaaca gaaatccaac 840 ctagagetge teegeatete eetgetgete ateeagtegt ggetggagee egtgeagtte 900 ctcaggagtg tcttcgccaa cagcctggtg tacggcgcct ctgacagcaa cgtctatgac 960 ctcctaaagg acctagagga aggcatccaa acgctgatgg ggaggctgga agatggcagc 1020 ccccggactg ggcagatctt caagcagacc tacagcaagt tcgacacaaa ctcacacaac 1080 gatgacgcac tactcaagaa ctacgggctg ctctactgct tcaggaagga catggacaag 1140 gtcgagacat tcctgcgcat cgtgcagtgc cgctctgtgg agggcagctg tggcttcggc 1200

```
<210> 74
<211> 412
<212>
       PRT
<213>
       Artificial
<220>
       synthetic sequence
<223>
<220>
<221> MISC_FEATURE
<222>
       (193)..(398)
       sequence is repeated N-1 times, where N is a positive whole numbe
<400> 74
Ser Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu
Arq Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
                                25
Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn
Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn
Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser
                                        75
Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser
Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
                               105
Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg
        115
                             120
Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr
Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn
                    150
                                        155
Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr
                                  170
                165
Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe
                                 185 -
```

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Phe

Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu 230 Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu 265 Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu Leu 280 Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe 295 Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys 345 Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly 360 Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu 375 380 Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly <210> 75 <211> 2445 <212> DNA <213> Artificial

<220>

synthetic sequence <223>

<220>

<221> misc_feature

(1237)..(1854) <222>

sequence is repeated N-1 times, where N is a positive whole numbe <223>

| catatgttcc | caaccattcc | cttatccagg | ctttttgaca | acgctatgct | ccgcgcccat | 60 |
|------------|------------|------------|------------|------------|------------|------|
| cgtctgcacc | agctggcctt | tgacacctac | caggagtttg | aagaagccta | tatcccaaag | 120 |
| gaacagaagt | attcattcct | gcagaacccc | cagacctccc | tctgtttctc | agagtctatt | 180 |
| ccgacaccct | ccaacaggga | ggaaacacaa | cagaaatcca | acctagagct | gctccgcatc | 240 |
| tecetgetge | tcatccagtc | gtggctggag | cccgtgcagt | tcctcaggag | tgtcttcgcc | 300 |
| aacagcctgg | tgtacggcgc | ctctgacagc | aacgtctatg | acctcctaaa | ggacctagag | 360 |
| gaaggcatcc | aaacgctgat | ggggaggctg | gaagatggca | gcccccggac | tgggcagatc | 420 |
| ttcaagcaga | cctacagcaa | gttcgacaca | aactcacaca | acgatgacgc | actactcaag | 480 |
| aactacgggc | tgctctactg | cttcaggaag | gacatggaca | aggtcgagac | attcctgcgc | 540 |
| atcgtgcagt | gccgcțctgt | ggagggcagc | tgtggcttcg | gcggcggcgg | atcaggcggc | 600 |
| ggcggatcag | geggeggegg | atcattccca | accattccct | tatccaggct | ttttgacaac | 660 |
| gctatgctcc | gcgcccatcg | tctgcaccag | ctggcctttg | acacctacca | ggagtttgaa | 720 |
| gaagcctata | tcccaaagga | acagaagtat | tcattcctgc | agaaccccca | gacctccctc | 780 |
| tgtttctcag | agtctattcc | gacaccctcc | aacagggagg | aaacacaaca | gaaatccaac | 840 |
| ctagagctgc | teegeatete | cctgctgctc | atccagtcgt | ggctggagcc | cgtgcagttc | 900 |
| ctcaggagtg | tcttcgccaa | cagcctggtg | tacggcgcct | ctgacagcaa | cgtctatgac | 960 |
| ctcctaaagg | acctagagga | aggcatccaa | acgctgatgg | ggaggctgga | agatggcagc | 1020 |
| ccccggactg | ggcagatctt | caagcagacc | tacagcaagt | tcgacacaaa | ctcacacaac | 1080 |
| gatgacgcac | tactcaagaa | ctacgggctg | ctctactgct | tcaggaagga | catggacaag | 1140 |
| gtcgagacat | tcctgcgcat | cgtgcagtgc | cgctctgtgg | agggcagctg | tggcttcggc | 1200 |
| ggcggcggat | caggcggcgg | cggatcaggc | ggcggcggat | cattcccaac | cattccctta | 1260 |
| tccaggcttt | ttgacaacgc | tatgctccgc | gcccațcgtc | tgcaccagct | ggcctttgac | 1320 |
| acctaccagg | agtttgaaga | agcctatatc | ccaaaggaac | agaagtattc | attcctgcag | 1380 |
| aacccccaga | cctccctctg | tttctcagag | tctattccga | caccctccaa | cagggaggaa | 1440 |
| acacaacaga | aatccaacct | agagctgctc | cgcatctccc | tgctgctcat | ccagtcgtgg | 1500 |
| ctggagcccg | tgcagttcct | caggagtgtc | ttcgccaaca | gcctggtgta | cggcgcctct | 1560 |
| gacagcaacg | tctatgacct | cctaaaggac | ctagaggaag | gcatccaaac | gctgatgggg | 1620 |
| aggctggaag | atggcagccc | ccggactggg | cagatettea | agcagaccta | cagcaagttc | 1680 |
| gacacaaact | cacacaacga | tgacgcacta | ctcaagaact | acgggctgċt | ctactgcttc | 1740 |

÷

```
aggaaggaca tggacaaggt cgagacattc ctgcgcatcg tgcagtgccg ctctgtggag
                                                                     1800
ggcagctgtg gcttcggcgg cggcggatca ggcggcggcg gatcaggcgg cggcggatca
                                                                     1860
ttcccaacca ttcccttatc caggettttt gacaacgeta tgctccgcgc ccatcgtctg
                                                                     1920
                                                                     1980
caccagctgg cctttgacac ctaccaggag tttgaagaag cctatatccc aaaggaacag
aagtatteat teetgeagaa eeeeeagaee teeetetgtt teteagagte tatteegaea
                                                                     2040
ccctccaaca gggaggaaac acaacagaaa tccaacctag agctgctccg catctccctg
                                                                     2100
etgeteatee agteqtqqct qqaqeeeqtg cagtteetea ggagtgtett eqceaacage
                                                                     2160
ctgqtgtacq gcgcctctga cagcaacgtc tatgacctcc taaaggacct agaggaaggc
                                                                     2220
atccaaacgc tgatggggag gctggaagat ggcagccccc ggactgggca gatcttcaag
                                                                     2280
caqacctaca qcaaqttcqa cacaaactca cacaacgatg acgcactact caagaactac
qqqctqctct actgcttcag gaaggacatg gacaaggtcg agacattcct gcgcatcgtg
                                                                     2400
cagtgccgct ctgtggaggg cagctgtggc ttctaggtcg acgcg
                                                                     2445
<210>
       76
<211>
       810
<212>
       PRT
<213>
       Artificial
<220>
<223>
       synthetic sequence
<220>
<221>
       MISC FEATURE
      (412)..(617)
<222>
       sequence is repeated N-1 times, where N is a positive whole numbe
<220>
<221>
       mat_peptide .
       (1)..()
<222>
<400> 76
Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu
Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn
```

Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn

| Arg 65 | Glu | Glu | Thr | Gln | Gln 70 | Lys | Ser | Asn | Leu | Glu 75 | Leu | Leu | Arg | Ile | Ser 80 |
|------------|------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Leu | Leu | Leu | Ile | Gln 85 | Ser | Trp | Leu | Glu | Pro 90. | Val | Gln | Phe | Leu | Arg 95 | Ser |
| Val | Phe | Ala | Asn 100 | Ser | Leu | Val | Tyr | Gly 105 | Ala | Ser | Asp | Ser | Asn 110 | Val | Tyr |
| Asp | Leu [.] | Leu 115 | Lys | Asp | Leu | Glu | Glu 120 | Gly | Ile | Gln | Thr | Leu 125 | Met | Gļy. | Arg |
| Leu | Glu 130 | | Gly | Ser | Pro | Arg 135 | Thr | Gly | Gln | Ile | Phe 140 | Lys | Gln | Thr | Tyr |
| Ser 145 | Lys | Phe | Asp | Thr | Asn 150 | Ser | His | Asn | Asp | Asp 155 | Ala | Leu | Leu | Lys | Asn 160 |
| Tyr | Gly | Leu | Leu | Tyr 165 | Cys | Phe | Arg | Lys | Asp 170 | Met | Asp | Lys | Val | Glu 175 | Thr |
| Phe | Leu | Arg | Ile 180 | Val | Gln | Cys | Arg | Ser 185 | Val | Glu | Gly | Ser | Cys 190 | Gly | Phe |
| Gly | Gly · | Gly 195 | Gly | Ser | Gly | Gly | Gly 200 | Gly | Ser | Gly | Gly | Gly 205 | Gly | Ser | Phe |
| Pro | Thr 210 | Ile | Pro | Leu | Ser | Arg 215 | Leu | Phe | Asp | Asn | Ala 220 | Met | Leu | Arg | Ala |
| 225 | | | | | 230 | | | | | 235 | | | | Glu | 240 |
| Ala | Tyr | Ile | Pro | Lys 245 | Glu | Gln | Lys | Tyr | Ser 250 | Phe | Leu | Gln | Asn | Pro 255 | Gln |
| | | | 260 | | | | | 265 | | | | | 270 | Arg | |
| | | 275 | | | | | 280 | | | | | 285 | | Leu | |
| | 290 | | | | | 295 | | | | | 300 | | | Val | |
| 305 | | | | | 310 | | | | | 315 | | | | Asp | 320 |
| | • | | | 325 | | | • | | 330 | | | | | Leu 335 | |
| | . • | | 340 | | | | | 345 | | | | | 350 | Ser | - |
| Phe | Asp | Thr | Asn | Ser | His. | Asn. | Asp | Asp | Ala | Len | Leu | Lvs | Asn | Tvr | Glv |

Leu Leu Tyr Cys Phe Arq Lys Asp Met Asp Lys Val Glu Thr Phe Leu 375

Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Gly Gly 395

Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Phe Pro Thr

Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg 425

Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr

Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser 455

Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr 470 475

Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu Leu Leu Ile 490

Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn 505

Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys

Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu Glu Asp Gly 535

Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp

Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu

Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile

Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Gly Gly Gly

Ser Gly Gly Gly Ser Gly Gly Gly Ser Phe Pro Thr Ile Pro 615 620

Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu His 625

Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro 650

Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys

660 665 670

Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr Gln Gln 680 Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu 710 Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln 790 795 Cys Arg Ser Val Glu Gly Ser Cys Gly Phe <210> 77 <211> 593 DNA <212> <213> Artificial <220> <223> synthetic sequence <400> catatgttcc caaccattcc cttatccagg ctttttgaca acgctatgct ccgcgcccat 60 cgtctgcacc agctggcctt tgacacctac caggagtttg aagaagccta tatcccaaag 120 gaacagaagt attcatteet geagaaceee cagaeeteee tetgtttete agagtetatt cegacaceet ceaacaggga ggaaacacaa cagaaateea acetagaget geteegeate . 240 tecetgetge teatecagte gtggetggag ceegtgeagt teeteaggag tgtettegee 300 aacageetgg tgtaeggege etetgaeage aaegtetatg aceteetaaa ggaeetagag 360 gaaggcatee aaacgetgat ggggaggètg gaagatggea geeeceggae tgggeagate 420 ttcaagcaga cctacagcaa gttcgacaca aactcacaca acgatgacgc actactcaag 480 aactacgggc tgctctactg cttcaggaag gacatggaca aggtcgagac attcctgcgc 540 ategtgeagt geegetetgt ggaqqqeage tqtqqettee atggateqaa tte 593

```
<210>
       78
<211>
       192
       PRT
<212>
<213>
      Artificial
<220>
<223>
      synthetic sequence
<400> 78
Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu
                                    10
Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
                                25
Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn
Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn
                     . 55
Arg Glu Glu Thr Gln Gln Lys Şer Asn Leu Glu Leu Leu Arg Ile Ser
                                        75
Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser
Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
                                105
Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg
        115
Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr
                        135
Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn
                    150
                                        155
Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr
Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe
                                185
<210>
       79
<211>
       592
<212>
      DNA
<213>
      Artificial
<220>
<223>
      synthetic sequence
```

<400> 79

| aagctttccc | aaccattccc | ttatccaggc | tttttgacaa | cgctatgctc | cgcgcccatc | 60 |
|------------|------------|------------|------------|------------|------------|-----|
| gtctgcacca | gctggccttt | gacacctacc | aggagtttga | agaagcctat | atcccaaagg | 120 |
| aacagaagta | ttcattcctg | cagaaccccc | agacctccct | ctgtttctca | gagtctattc | 180 |
| cgacaccctc | caacagggag | gaaacacaac | agaaatccaa | cctagagctg | ctccgcatct | 240 |
| ccctgctgct | catccagtcg | tggctggagc | ccgtgcagtt | cctcaggagt | gtcttcgcca | 300 |
| acagcctggt | gtacggcgcc | tctgacagca | acgtctatga | cctcctaaag | gacctagagg | 360 |
| aaggcatcca | aacgctgatg | gggaggctgg | aagatggcag | cccccggact | gggcagatct | 420 |
| tcaagcagac | ctacagcaag | ttcgacacaa | actcacacaa | cgatgacgca | ctactcaaga | 480 |
| actaçgggçt | gctctactgc | ttcaggaagg | acatggacaa | ggtcgagaca | ttcctgcgca | 540 |
| tcgtgcagtg | ccgctctgtg | gagggcagct | gtggcttcca | tggatcgaat | tc | 592 |

<210> 80

<211> 191

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<400> 80

Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg 1 5 10 15

Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu 20 25 30

Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro 35 40 45

Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg
50 55 60

Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu 65 70 75 80

Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val 85 90 95

Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp 100 105 110

Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu
115 120 125

Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser 130 135 140 Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr 150 145 Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe <210> 81 <211> 587 <212> DNA <213> Artificial <220> <223> synthetic sequence <400> aagettteee aaceatteee ttatecagge tttttgacaa egetatgete egegeecate 60 gtctgcacca gctggccttt gacacctacc aggagtttga agaagcctat atcccaaagg 120 aacagaagta ttcattcctg cagaaccccc agacctccct ctgtttctca gagtctattc 180 cgacaccete caacagggag gaaacacaac.agaaatecaa.cetagagetg etcegcatet 240 ccctgctgct catccagtcg tggctggagc ccgtgcagtt cctcaggagt gtcttcgcca 300 acagcetggt gtacggcgcc tetgacagca acgtetatga cetectaaag gacetagagg 360 aaggcatcca aacgctgatg gggaggctgg aagatggcag cccccggact gggcagatct 420 tcaagcagac ctacagcaag ttcgacacaa actcacacaa cgatgacgca ctactcaaga, 480 actacgggct gctctactgc ttcaggaagg acatggacaa ggtcgagaca ttcctgcgca 540 tcgtgcagtg ccgctctgtg gagggcagct gtggcttcta gggatcc 587 <210> 82 <211> 191 <212> PRT <213> Artificial <220> synthetic sequence <223> <400> 82 Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu 25

Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro

35

Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arq Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu 70 75 Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp 105 Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu 115 Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser 135 Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr 145 150 155 Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe 185 <210> 83 <211> 1165 <212> DNA <213> Artificial <220> <223> synthetic sequence <220> <221> misc_feature <222> (579)..(1151) <223> sequence is repeated N-1 times, where N is a positive whole numbe <400> 83 aagettteee aaccatteee ttatecagge tttttgacaa egetatgete egeqeecate 60 gtctgcacca gctggccttt gacacctacc aggagtttga agaagcctat atcccaaagg 120 aacagaagta ttcattcctg cagaaccccc agacctccct ctgtttctca gagtctattc 180 cgacaccete caacagggag gaaacacaac agaaatecaa cetagagetg etecgeatet 240 ecctgetget catecagteg tggetggage cegtgeagtt ceteaggagt gtettegeea 300 acageetggt gtaeggegee tetgaeagea aegtetatga eeteetaaag gaeetagagg 360

| aaggcatcca | aacgctgatg | gggaggctgg | aagatggcag | cccccggact | gggcagatct | 420 |
|------------|------------|------------|------------|------------|------------|------|
| tcaagcagac | ctacagcaag | ttcgacacaa | actcacaçaa | cgatgacgca | ctactcaaga | 480 |
| actacgggct | gctctactgc | ttcaggaagg | acatggacaa | ggtcgagaca | ttcctgcgca | 540 |
| tcgtgcagtg | ccgctctgtg | gagggcagct | gtggcttctt | cccaaccatt | cccttatcca | 600 |
| ggctttttga | caacgctatg | ctccgcgccc | atcgtctgca | ccagctggcc | tttgacacct | 660 |
| accaggagtt | tgaagaagcc | tatatcccaa | aggaacagaa | gtattcattc | ctgcagaacc | 720 |
| cccagacctc | cctctgtttc | tcagagtcta | ttccgacacc | ctccaacagg | gaggaaacac | 780 |
| aacagaaatc | caacctagag | ctgctccgca | tctccctgct | gctcatccag | tcgtggctgg | 840 |
| agcccgtgca | gttcctcagg | agtgtcttcg | ccaacagcct | ggtgtacggc | gcctctgaca | 900 |
| gcaacgtcta | tgacctccta | aaggacctag | aggaaggcat | ccaaacgctg | atggggaggc | 960 |
| tggaagatgg | cagcccccgg | actgggcaga | tcttcaagca | gacctacagc | aagttcgaca | 1020 |
| caaactcaca | caacgatgac | gcactactca | agaactacgg | gctgctctac | tgcttcagga | 1080 |
| aggacatgga | caaggtcgag | acattcctgc | gcatcgtgca | gtgccgctct | gtggagggca | 1140 |
| gctgtggctt | ccatggatcg | aattc | . • | | | 1165 |

<210> 84

<211> 191

<212> PRT

<213> Artificial

<220>

<223> synthetic sequence

<220>

<221> MISC_FEATURE

<222> (1)..(191)

<223> sequence is repeated N times, where N is a positive whole number

<400> 84

Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg 1 5 10 15

Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu 20 25 30

Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg 50 55 60

| G1u 65 | Glu | Thr | GIn | GIn | Lys 70 | Şer | Asn | Leu | Glu | Leu 75 | Leu | Arg | Ile | Ser | Leu 80 | |
|------------------------------|--------------|----------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------|
| Leu | Leu | Ile | Gln | Ser 85 | Trp | Leu | Glu | Pro | Val 90 | Gln | Phe | Leu | Arg | Ser 95 | Val | |
| Phe | Aļa | Asn | Ser 100 | Leu | Val | Tyr | Gly | Ala 105 | Ser | Asp | Ser | Asn | Val 110 | Tyr | Asp | |
| Leu | Leu | Lys 115 | Asp | Leu | Glu | Glu | Gly 120 | Ile | Gln | Thr | Leu | Met 125 | Gly | Arg | Leu | |
| Glu | Asp 130 | Gly | Ser | Pro | Arg | Thr 135 | Gly | Gln | Ile | Phe | Lys 140 | Gln | Thr | Tyr | Ser | |
| Lys 145 | Phe | Asp | Thr | Asn | Ser 150 | His | Asn | Asp | Asp | Ala 155 | Leu | Leu | Lys | Asn | Tyr 160 | . 3 |
| Gly | Leu | Leu | Tyr | Cys 165 | Phe | Arg | Lys | Asp | Met 170 | Asp | Lys | Val | Glu | Thr 175 | Phe | |
| Leu | Arg | Ile | Val 180 | Gln | Cys | Arg | Ser | Val 185 | Glu | Gly | Ser | Cys | Gly 190 | Phe | | |
| <210 <211 <211 <211 | l> 2 2> I | 35 2307 DNA Artii | Eicia | al | | | | | | | | | | | | - |
| <220 <220 | | syntl | netio | c sec | quen | ce | | | | | | | | | | |
| <220 <221 <222 <221 | L> r 2> | (115 | | (172 | | ated | N-1 | time | es, 1 | where | e N : | is a | pos | itiv | e whole | e numbe |
| <400 | | 35 | | | | | | | | | | | | | | |
| | | | caaco | catto | cc c1 | tato | ccag | g cti | tttt | gaca | acg | ctaț | gct | ccgc | gcccat | 60 |
| cgt | ctgca | acc a | agct | ggcct | tt to | gaca | ccta | c cag | ggagt | tttg | aaga | aagc | cta | tatco | ccaaag | 120 |
| gaad | cagaa | agt a | attca | attco | ct go | cagaa | accc | c ca | gacci | tccc | tct | gttt | ctc | agagt | tctatt | 180 |
| ççga | acaco | cct o | ccaa | caggg | ga gg | gaaad | cacaa | a caq | gaaal | cca | acci | tagag | gct | gctc | cgcatc | 240 |
| tcc | ctgct | igc t | cato | ccagt | tc gi | ggct | tggag | g cc | cgtg | cagt | tcc | tcag | gag | tgtc | ttegee | 300 |
| aaca | agcct | egg t | gtad | cggcg | gc ct | ctga | acago | c aad | cgtc | tatg | acc | tecta | aaa | ggac | ctagag | 360 |
| gaag | ggcat | ccc a | aaac | gctga | at g | gggag | ggctg | g gaa | agato | ggca | gcc | ccg | gac | tggg | cagato | 420 |
| ttca | aagca | aga d | ccta | cagca | aa gi | tcga | acaca | a aad | ctca | çaça | acga | atga | ege | açta | ctcaag | 480 |

, *č*

| aactacgggc | tgctctactg | cttcaggaag | gacatggaca | aggtcgagac | attcctgcgc | 540 |
|------------|------------|------------|------------|------------|------------|-------|
| atcgtgcagt | gccgctctgt | ggagggcagc | tgtggcttct | tcccaaccat | tcccttatcc | 600 |
| aggctttttg | acaacgctat | gctccgegcc | categtetge | accagetgge | ctttgacacc | 660 |
| taccaggagt | ttgaagaagc | ctatatecea | aaggaacaga | agtattcatt | cctgcagaac | 720 |
| ccccagacct | ccctctgttt | ctcagagtct | attccgacac | cctccaacag | ggaggaaaca | 780 |
| caacagaaat | ccaacctaga | gctgctccgc | atctccctgc | tgctcatcca | gtcgtggctg | 8.40 |
| gagcccgtgc | agttcctcag | gagtgtcttc | gccaacagcc | tggtgtacgg | cgcctctgac | 900 |
| agcaacgtct | atgacctcct | aaaggaccta | gaggaaggca | tccaaacgct | gatggggagg | : 960 |
| ctggaagatg | gcagcccccg | gactgggcag | atcttcaagc | agacctacag | caagttcgac | 1020 |
| acaaactcac | acaacgatga | cgcactactc | aagaactacg | ggctgctcta | ctgcttcagg | 1080 |
| aaggacatgg | acaaggtcga | gacattcctg | cgcatcgtgc | agtgccgctc | tgtggagggc | 1140 |
| agctgtggct | tcttcccaac | cattccctta | tccaggcttt | ttgacaacgc | tatgctccgc | 1200 |
| gcccatcgtc | tgcaccagct | ggcctttgac | acctaccagg | agtttgaaga | agcctatatc | 1260 |
| ccaaaggaac | agaagtattc | attcctgcag | aacccccaga | cctccctctg | tttctcagag | 1320 |
| tctattccga | caccctccaa | cagggaggaa | acacaacaga | aatccaacct | agagctgctc | 1380 |
| cgcatctccc | tgctgctcat | ccagtcgtgg | ctggagcccg | tgcagttcct | caggagtgtc | 1440 |
| ttcgccaaca | gcctggtgta | cggcgcctct | gacagcaacg | tctatgacct | cctaaaggac | 1500 |
| ctagaggaag | gcatccaaac | gctgatgggg | aggctggaag | atggcagccc | ccggactggg | 1560 |
| cagatettea | agcagaccta | cagcaagttc | gacacaaact | cacacaacga | tgacgcacta | 1620 |
| ctcaagaact | acgggctgct | ctactgcttc | aggaaggaca | tggacaaggt | cgagacattc | 1680 |
| ctgcgcatcg | tgcagtgccg | ctctgtggag | ggcagctgtg | gcttcttccc | aaccattccc | 1740 |
| ttatccaggc | tttttgacaa | cgctatgctc | cgcgcccatc | gtctgcacca | gctggccttt | 1800 |
| gacacctacc | aggagtttga | agaagcctat | atcccaaagg | aacagaagta | ttcattcctg | 1860 |
| cagaaccccc | agacctccct | ctgtttctca | gagtctattc | cgacaccctc | caacagggag | 1920 |
| gaaacacaac | agaaatccaa | cctagagctg | ctccgcatct | ccctgctgct | catccagtcg | 1980 |
| tggctggagc | ccgtgcagtt | cctcaggagt | gtettegeca | acagcctggt | gtacggcgcc | 2040 |
| tctgacagca | acgtctatga | cctcctaaag | gacctagagg | aaggcatcca | aacgctgatg | 2100 |
| gggaggctgg | aagatggcag | ccccggact | gggcagatct | tcaagcagac | ctacagcaag | 2160 |
| ttcgacacaa | actcacacaa | cgatgacgca | ctactcaaga | actacgggct | gctctactgc | 2220 |

```
ttcaqgaagg acatggacaa ggtcgagaca ttcctgcgca tcgtgcagtg ccgctctgtg
gagggcagct gtggcttcta gggatcc
                                                                   2307
<210> 86
<211>
      192
<212> PRT
<213> Artificial
<220>
<223> synthetic sequence
<220>
<221> MISC_FEATURE
<222> (2)..(192)
<223> sequence is repeated N+2 times, where N is a positive whole numbe
<220>
<221> mat peptide
      (1)..()
<222>
<400> 86
Met Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu
      . 5
Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe
                               25
Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn
Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn
Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser
                   70
Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser
Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr
Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg
                           120
       115
Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr
                       135
Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn
                   150
                               155
```

Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr 165 170 175

Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe 180 185 190